UNIVERSITÀ DEGLI STUDI DI SIENA

Facoltà di Scienze Matematiche Fisiche e Naturali

Tesi di dottorato in Fisica Sperimentale
PhD Thesis in Experimental Physics
XXIV Ciclo

Measurements of branching fraction ratios and CP-asymmetries in suppressed $B^- \to D(\to K^+\pi^-)K^-$ and $B^- \to D(\to K^+\pi^-)\pi^-$ decays

Supervisor

Prof. Giovanni Punzi

Tutor

Doct. Paola Squillacioti

Candidate

Paola Garosi

Contents

Contents

Introduction

Since 1964, year of the discovery of CP violation in the kaon system, CP violation has played a central role in particle physics. The large CP-violating asymmetries predicted, and later observed, in the B meson system confirmed the CKM picture of the phenomenon, but opened further questions and stimulated an increase of precision of the measurements in the search for new physics sources of CP violation.

The "cosmological question" of the dominance of matter over antimatter in our universe is an important hint for the presence of CP-violating effects of a much larger strength than the standard CKM picture can accomodate.

Experimental tests of the CKM picture can be done by measuring the angles and the sides of the so called "Unitarity Triangle", using a variety of neutral B mixing and neutral and charged B decays. Today a good level of precision has been reached for two of the angles, but the resolution on the third (γ) angle is still rather poor, and limited by the size of the available B samples. A precise measurement of γ in "theoretically clean" decays where it appears at tree-level (most notably the family of $B^- \to DK^-$ decay modes), is important not only as a fundamental parameter, but also as a reference point for decays where the presence of significant loop contribution may exhibit additional CP-violating effects due to BSM processes.

The large production of B species at hadron colliders can in principle be exploited to improve the precision of our knowledge of the angle γ. However, the rarity of the interesting decay modes and the presence of much larger backgrounds compared to e^+e^- colliders presents a significant challenge. In this Thesis we present the first successful reconstruction of the rare doubly-suppressed $B^- \to D(\to K^+\pi^-)h^-$ decays ($h = K, \pi$) in hadron collisions, and the measurement of their branching

fractions and CP asymmetries for the extraction of the angle γ [1, 2]. The analysis was performed on a 7 fb^{-1} sample of data collected by the CDF experiment at the Tevatron collider, using a specialized trigger on displaced tracks.

The main challenge of the analysis is the presence of similar backgrounds in large quantities, most notably the $B^- \to D(\to K^- \pi^+)h^-$, which are two (in the case of the kaon mode) or three (for the pion mode) orders of magnitude more favored then the signal decays. We disentangle the signal and background contributions by an unbinned maximum likelihood fit, combining information on the kinematics of the decay with information on particle identification of the final state particles coming from their specific energy loss in crossing the detector. This requires a careful determination of the features of all backgrounds and precision in-situ calibration of the particle identification observables. The similarity of the physics backgrounds is exploited as an advantage of the analysis by using them as calibration/reference modes for the signals being searched, leading to a fully data-driven analysis approach.

As a result, we are able to obtain statistically significant signals for the suppressed $B^- \to D\pi^-$ and $B^- \to DK^-$ decays, and find branching fraction ratio and CP-asymmetries in agreement with results from other experiment.

This Thesis is organized as follows. In Chapter 1 we give the theoretical motivation for studying CP violation and for the measurement of the γ angle. In Chapter 2 we describe the Tevatron collider and the CDF experiment, giving more resonance to the sub-detectors used in the current measurement. Chapter 3 describes the selection of the favored and suppressed samples and the optimization of the criteria to reduce the backgrounds. Chapter 4 describes the background study and the implementation of the maximum likelihood fit. In the Chapter 5 fit results are extracted, checked to be intrinsically consistent and corrected for detector acceptance. In Chapter 6 we give the systematic uncertainties evaluation and finally in Chapter 7 final results are reported and compared with measurements from other experiments.

1

CP violation and methods for the measurement of the γ angle

In this Chapter we will describe the theoretical motivations for studying CP violation and for measuring the γ angle. These are the basis for the measurement described in this Thesis.

1.1 The standard model

The idea that the reality can be made by simple constituents was present also in ancient times. Only in the past two centuries, with the development of new technologies, we could discover that this simple idea can be true.

Nowadays the Standard Model (SM) [3, 4, 5] is the physics theory describing in the most accurate and complete way those fundamental elements and their interactions. It predicts the existence of four interaction forces, the *electomagnetic*, the *weak*, the *strong*, and the *gravitational* interaction [6, 7]. Using the quantum field theory and the special relativity, the SM gives a comprehensive model for the unification of these interactions, with the exclusion of the gravitational force. The strain to unify the models is made to obtain the simplest and more elegant theory to describe the reality. The SM postulates that all the interactions are direct consequence of a charge conservation under a *local gauge transformation* of the free

fields: to leave the Lagrangian symmetric with respect to the gauge transformations, we need to introduce the "gauge bosons", mediators of the interactions: the *photon*, which mediates the electromagnetic interaction among charged particles; the $W^{\pm}$ *and* Z^0 *bosons*, which mediate the weak interaction among flavored particles, and the *gluons*, which mediate the strong interaction among colored particles.

The symmetry group of the standard model is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where the first term is the irreducible group of the symmetry of the strong interaction, the second of the weak interaction, and the third of the electromagnetic interaction [8, 9]. The elementary particles subjected to these interactions are massless. Mass is mediated by an other boson, the Higgs boson, which comes from the spontaneous breaking of the symmetry. This boson has not yet been seen experimentally.

The elementary particles are divided into *leptons* and *quarks*. Both of them are fermions and are subjected to the electroweak interaction. Quarks are also affected by strong interaction, being the fundamental constituents of the hadrons. They are all divided into three families. For each family we have a $SU(2)$ left-handed doublet (L), that couples with the weak interaction, and a right-handed singlet (R), not affected by weak interaction.

For the leptons we have:

$$\psi_L = \begin{pmatrix} \nu_i \\ \ell_i^- \end{pmatrix}_L \qquad \psi_R = \ell_R^-$$

where i runs over the three lepton families (*electrons*, *muons*, and *taus*), ν is the neutral massless[1] neutrino and ℓ is the charged massive lepton.

For the quarks we have:

$$\psi_L = \begin{pmatrix} u_i \\ d_i' \end{pmatrix}_L$$

and two singlets u_{iR} and d_{iR}.

$u_i = (u, c, t)$ are the three quarks "up"-like, with charge 2/3 e (where e is the elementary charge), and $d_i' = (d', s', b')$ are the weak interaction eigenstates, with charge $-1/3$ e. The weak eigenstates are obtained rotating the mass eigenstates d_i by the coefficients V_{ij} given by the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix (see

---

[1]We are considering only the SM formalism, which predicts massless neutrinos.

Cap. 1.2), $d'_i \equiv \sum_{ij} V_{ij} d_j$.

The three families of quarks are (u, d), (c, s) and (t, b), namely (up, down), (charm, strange) and (top, bottom (or beauty)).

Symmetries are of fundamental importance in the SM, and they can be continuous or discrete. In particular the discrete symmetries are: *parity* (P) that reverses the spatial coordinates, *charge* (C) that exchanges the quantum numbers of the particles with the antiparticles, and *time* (T), that reverses the time coordinates [10, 11].

All interactions are invariant under those symmetries, except the weak interaction that violates both C and P. In fact, for example, only neutrino left-handed ν_L and anti-neutrino right-handed $\bar{\nu}_R$ exist. Applying C to those states we obtain two particles that does not exist in nature: $\text{C}|\nu_L\rangle = |\bar{\nu}_L\rangle$ and $\text{C}|\bar{\nu}_R\rangle = |\nu_R\rangle$. This is the same for parity: $\text{P}|\nu_L\rangle = |\nu_R\rangle$ and $\text{P}|\bar{\nu}_R\rangle = |\bar{\nu}_L\rangle$.

When there is the simultaneous application of both C and P symmetry, the weak interaction is conserved most of the time. In the example of the neutrino we have: $\text{CP}|\nu_L\rangle = |\bar{\nu}_R\rangle$ and $\text{CP}|\bar{\nu}_R\rangle = |\nu_L\rangle$. In these cases the system is invariant under CP and its Lagrangian $\mathcal{L}$ is real.

There are some cases when the weak interaction can violate CP. This arises mathematically when the Lagrangian of the system has complex elements.

1.2 The CKM matrix

In the SM the CP-symmetry is violated by a complex phase appearing in the Yukawa coupling of the Higgs boson to the quarks. As a consequence, the $W^\pm$ boson couples with the physics quarks u_{iL} and d_{iL} through the coefficients given by the unitarity CKM matrix [12, 13]:

$$V \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

As a consequence, charged weak interactions can change the flavor of a quark. For example, a u quark can couple not only with a d quark, but also with a s or b. Flavor changing neutral currents are instead forbidden for tree-level amplitudes.

1. CP violation and methods for the measurement of the γ angle

Since the physical observables does not change by a phase redefinition of the quark fields, we want to verify that there is an irreducible complex phase.

For a generic $n \times n$ complex matrix we have $2n^2$ real parameters, while the unitarity condition $\sum_j V_{ij} V_{jk}^* = \delta_{ik}$ reduces them to n^2. We can redefining the $2n$ phases of the quark fields, and since the overall phase is irrelevant, $2n - 1$ relative phases can be removed from V. Moreover, since a unitarity matrix is also orthogonal, there are $\frac{1}{2} n(n - 1)$ parameters corresponding to independent rotation angles. Therefore the number of left independent phases in V is $\frac{1}{2}(n - 1)(n - 2)$.

In the case of two families of quarks, the number of angles is 1 (corresponding to the "Cabibbo angle") and the phases are 0 (no CP violation). In case of three families of quarks we have 3 angles and 1 irreducible phase, the one giving CP violation.

In conclusion, to have CP violation we need at least three families of non degenerate quarks (in case of degeneracy we can introduce an arbitrary rotation that can remove the complex phase). This mechanism was postulated when only three of the six quarks were known and it encouraged the search of the other three quarks.

There are several representation of the CKM matrix, the "standard" representation [14] is

$$
V = \begin{pmatrix}
c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\
-s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\
s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13}
\end{pmatrix},
$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ for the Euler angles θ_{ij} (i, j are the family labels), and δ is the CP-violating phase. Conventionally we can choose θ_{ij} in the first quadrant, so that $s_{ij}, c_{ij} \geq 0$.

Experimentally, matrix elements of V have different magnitudes. For example, there is a preference of the b quark to decay into c rather then u quarks. More generally $s_{13} \ll s_{23} \ll s_{12} \ll 1$.

The parameterization which emphasizes the order of magnitude of the elements was made by Wolfenstein [15]. We need to introduce four real parameters λ, A, ρ, and η, and, expanding in series of λ (0.2253 ± 0.0007 [16]), the CKM matrix can be written

as:

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta + i\eta\lambda^2/2) \\ -\lambda & 1 - \lambda^2/2 - i\eta A^2\lambda^4 & A\lambda^2(1 + i\eta\lambda^2) \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \ .$$

The hierarchy of the structure is manifest: up-type quarks prefer to couple with down-type quarks of the same family. The more two families are distant, the more the couplings are small.

1.3 The unitarity triangle

The unitarity condition of the CKM matrix implies that: $VV^+ = V^+V = 1$, or $\sum_k V_{ki}V_{kj}^* = \delta_{ij}$, where $k = \{u, c, t\}$, and $i, j = \{d, s, t\}$, with $i \neq j$. From these conditions we have twelve equations, half for the normalization and half for the orthogonality. The six orthogonality equations are:

$$
\begin{aligned}
ds: \quad & \underbrace{V_{ud}V_{us}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V_{ts}^*}_{\mathcal{O}(\lambda^5)} = 0 \\
uc: \quad & \underbrace{V_{ud}V_{cd}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{us}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{ub}V_{cb}^*}_{\mathcal{O}(\lambda^5)} = 0 \\
sb: \quad & \underbrace{V_{us}V_{ub}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{cb}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{ts}V_{tb}^*}_{\mathcal{O}(\lambda^2)} = 0 \\
ct: \quad & \underbrace{V_{cd}V_{td}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{ts}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{cb}V_{tb}^*}_{\mathcal{O}(\lambda^2)} = 0 \\
ut: \quad & \underbrace{V_{ud}V_{td}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{us}V_{ts}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{ub}V_{tb}^*}_{\mathcal{O}(\lambda^3)} = 0 \\
db: \quad & \underbrace{V_{ud}V_{ub}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{cd}V_{cb}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{td}V_{tb}^*}_{\mathcal{O}(\lambda^3)} = 0
\end{aligned}
$$

$\boxed{1.1}$

where we reported also the order of magnitude in powers of λ.

Since the matrix elements are complex, these six relations can be graphically represented as triangles in the complex plane.

The six triangles have very different shapes (Fig. 1.1), but they have the same area, half of the Jarlskog invariant J [17], which is a measure of CP violation, independent from the phase convention. It is defined by $J \equiv |Im[V_{ij}V_{kl}V_{il}^*V_{kj}^*]|$ and

$Im[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J\sum_{m,n}\epsilon_{ikl}\epsilon_{jln}$. J is defined as the only CP violation term in the CKM matrix: for this reason any CP-violating quantities in the SM must be proportional to J.
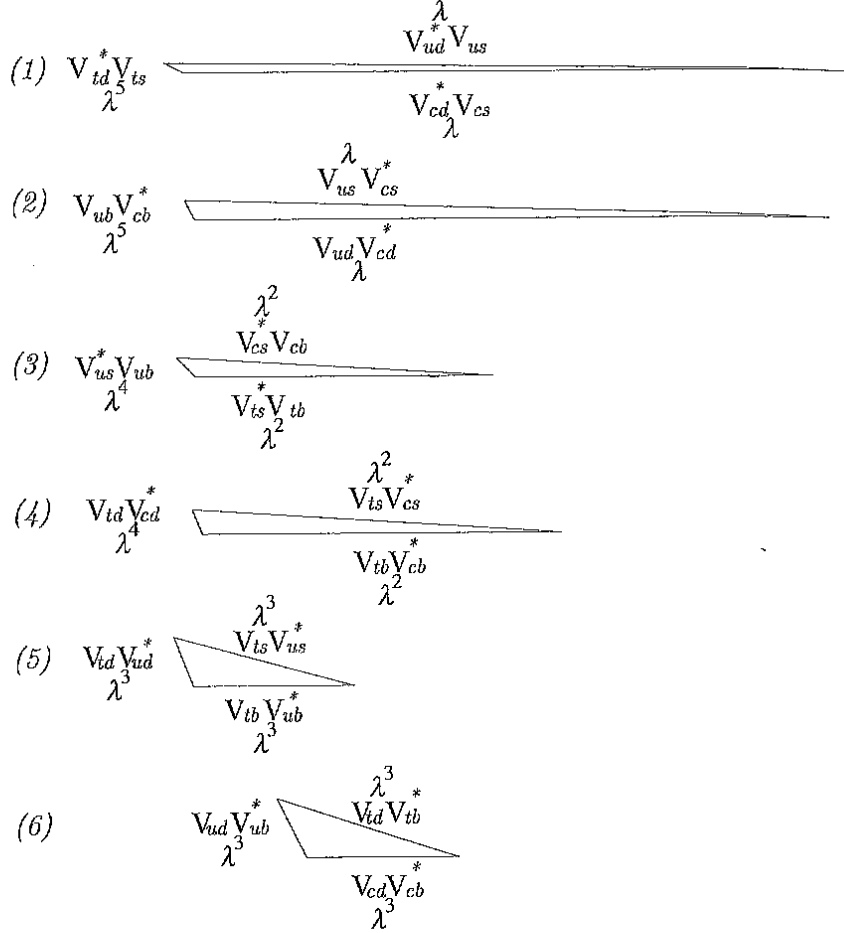
Figure 1.1: The six unitarity relations (eq. 1.1) represented as triangles in the complex plane.

The first relation in (1.1) corresponds to CP violation in K decays, the second to CP violation in D decays. The corresponding triangles are very squashed (even if they have the same area of the others), so that one of the angles representing the relative phases of the CKM elements is tiny.

Elements in the latter expression are those involved in B meson processes. Since the sides of the corresponding triangle are of the same order of magnitude, we expect to have large weak relative phases.

This is an important point for studying CP violation in B meson system: some decays of the B can have very large CP-violating asymmetry.

The latter triangle is often called "the unitarity triangle". We choose the phase convention in which $V_{cd}V_{cb}^*$ is real, and normalizing each side for its magnitude, we get a triangle of vertices in $(0,0)$, $(1,0)$, and (ρ,η) (Fig. 1.2) [18, 11].
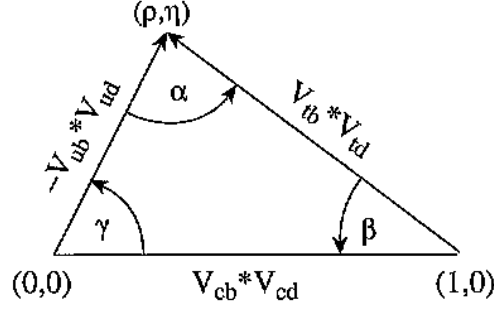
Figure 1.2: The unitarity triangle in the complex plane.

The two complex sides have lengths:

$$R_u \equiv \sqrt{\rho^2 + \eta^2} = \frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right| \qquad R_t \equiv \sqrt{(1-\rho)^2 + \eta^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \qquad \boxed{1.2}$$

The three angles, called α, β, and γ, are defined as:

$$
\begin{aligned}
\alpha &\equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \\
\beta &\equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \\
\gamma &\equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)
\end{aligned}
\qquad \boxed{1.3}
$$

(with $\alpha + \beta + \gamma = \pi$). The angle γ is in good approximation equal with the δ phase of the "standard" representation of the CKM matrix.

We can measure the sides and the angles of the unitarity triangle and verify the accuracy of the SM, as well as to probe new physics scenarios [19, 20, 21].

1.4 CP **violation in the** B **mesons system**

CP violation can occur in the sectors of the K, D, and B mesons, but we are going to give more details for the B sector, since larger CP-violating asymmetry can be

found. The formalism [9, 10] we are going to introduce works for all the three mesons.

The decay amplitude of the B meson (or of its CP conjugate $\bar{B}$) in a final state f ($\bar{f}$), can be written as:

$$A_f = \langle f|H|B\rangle \qquad \bar{A}_f = \langle f|H|\bar{B}\rangle$$
$$A_{\bar{f}} = \langle \bar{f}|H|B\rangle \qquad \bar{A}_{\bar{f}} = \langle \bar{f}|H|\bar{B}\rangle$$

where H is the total hamiltonian of the system.

The CP operator introduces a weak phase ϕ in the states:

$$CP|B\rangle = e^{i\phi_B}|\bar{B}\rangle \qquad CP|f\rangle = e^{i\phi_f}|\bar{f}\rangle$$
$$CP|\bar{B}\rangle = e^{-i\phi_B}|B\rangle \qquad CP|\bar{f}\rangle = e^{-i\phi_f}|f\rangle.$$

(1.4)

If CP is conserved A_f and $\bar{A}_{\bar{f}}$ have the same amplitude.

We can observe CP violation in three phenomena:

- in the decay;

- in the mixing;

- in the interference of decay and mixing.

1.4.1 CP violation in the decay

This process can happen for both charged and neutral mesons and it is called "direct CP violation". Summing over all contributions, we can write the decay amplitudes as:

$$A_f = \langle f|\mathcal{H}|B\rangle = \sum_j |A_j| e^{i(\delta_j + \phi_j)}$$

$$\bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{B}\rangle = \sum_j |A_j| e^{i(\delta_j - \phi_j)}$$

where δ is now the "strong CP-conserving phase", ϕ is the "weak phase" introduced in Eq. (1.4), and $\mathcal{H}$ is the hamiltonian of the weak interaction.

CP violation occurs when $|A_f|^2 \neq |\bar{A}_{\bar{f}}|^2$. This happens when there are at least two processes with different weak or strong phases. In this way the interfering term $|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -2\sum_{i,j} |A_i||A_j| \sin(\delta_i - \delta_j) \sin(\phi_i - \phi_j)$ is different from zero.

We define the CP-asymmetry, for both charged and neutral mesons, as:

$$\mathcal{A}_{CP} = \frac{\Gamma(B \to f) - \Gamma(\bar{B} \to \bar{f})}{\Gamma(B \to f) + \Gamma(\bar{B} \to \bar{f})} = \frac{\frac{|\bar{A}_{\bar{f}}|^2}{|A_f|^2} - 1}{\frac{|\bar{A}_{\bar{f}}|^2}{|A_f|^2} + 1},$$

where Γ is the decay rate. When the asymmetry is different from zero we have violation of CP.

For charged mesons decaying in two body, we have:

$$A_f = |a_1|e^{i(\delta_1 + \phi_1)} + |a_2|e^{i(\delta_2 + \phi_2)},$$

$$\bar{A}_{\bar{f}} = |a_1|e^{i(\delta_1 - \phi_1)} + |a_2|e^{i(\delta_2 - \phi_2)},$$

and the asymmetry becomes:

$$\mathcal{A}_{CP} = -\frac{2|a_1 a_2| \sin(\delta_2 - \delta_1) \sin(\phi_2 - \phi_1)}{|a_1|^2 + |a_2|^2 + 2|a_1 a_2| \cos(\delta_2 - \delta_1) \cos(\phi_2 - \phi_1)}, \qquad (1.5)$$

from which we can extract $(\phi_2 - \phi_1)$, knowing $\left|\dfrac{a_2}{a_1}\right|$ and $(\delta_2 - \delta_1)$.

1.4.2 CP **violation in the mixing**

The mixing, the oscillation between particles and anti-particles, can occur only in the neutral mesons. As an example, we can see in Fig. 1.3 the diagram of the mixing of the $B_s - \overline{B}_s$ system.

We can write the state at the time $t = 0$ as a superposition of the states B^0 and $\overline{B}^0$:

$$|\psi(0)\rangle = a(0)|B^0\rangle + b(0)|\overline{B}^0\rangle \,,$$

which evolves in time according to the Schroedinger equation, through the effective non-hermitian hamiltonian H [2], sum of a hermitian and anti-hermitian terms, respectively M and Γ. M is called the "mass matrix" and Γ is the "decay matrix".

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \equiv (M - \frac{i}{2}\Gamma)|\psi(t)\rangle \,.$$

CPT invariance ensures that:

$$\begin{aligned} M_{11} &= M_{22} = M_0 \\ \Gamma_{11} &= \Gamma_{22} = \Gamma_0 \,. \end{aligned}$$

---

[2]We can introduce this effective hamiltonian since we are not interested to the single decay products. Its non-hermitianity allows the states to decay and not only to oscillate.
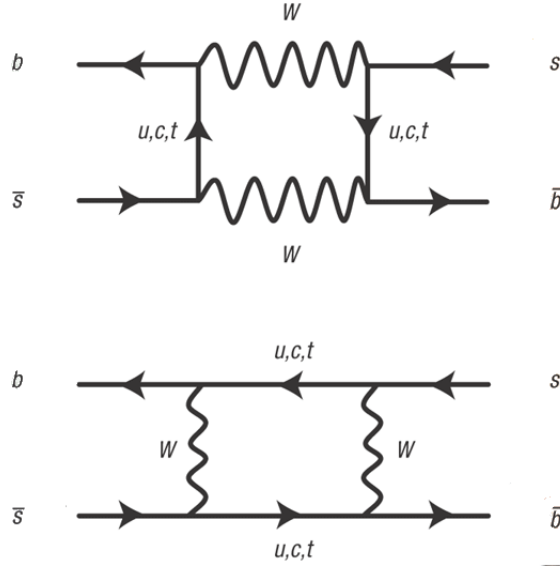
Figure 1.3: Diagram of the $B_s - \overline{B}_s$ system mixing.

Diagonalizing the system we can get the mass eigenstates, being L the lightest and H the eaviest:

$$
\begin{aligned}
|B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle \\
|B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle \, .
\end{aligned}
$$

These conditions are true:

- $|p|^2 + |q|^2 = 1$ (normalization)

- $\left(\dfrac{q}{p}\right)^2 = \dfrac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}} \, .$

and CP violation occurs when $\left|\dfrac{q}{p}\right| \neq 1$.

1.4.3 CP violation in the interference between decay and mixing

Consider the case when both B^0 and $\overline{B}^0$ decay into the same final state f, which is also CP eigenstate, $f = \overline{f}$:

$$
\begin{aligned}
B^0 &\to f \\
B^0 &\to \overline{B}^0 \to f
\end{aligned}
$$

and A_f and $\bar{A}_f$ being the corresponding decay amplitudes.

If we define $\lambda_f = \dfrac{q}{p}\dfrac{\bar{A}_f}{A_f}$ and $\Delta m_B = M_{B_H} - M_{B_L}$ (see Cap. 1.4.2), the time-dependent CP asymmetry can be written as:

$$
\begin{aligned}
\mathcal{A}_{CP}(t) &= \frac{\Gamma(B \to f)(t) - \Gamma(\overline{B} \to f)(t)}{\Gamma(B \to f)(t) + \Gamma(\overline{B} \to f)(t)} = \\
&= \left(\frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}\right)\cos\Delta m_B t - \left(\frac{2\Im m(\lambda_f)}{1 + |\lambda_f|^2}\right)\sin\Delta m_B t \ .
\end{aligned}
\qquad \boxed{1.6}
$$

The first term comes from the direct CP violation and it is zero when $|\lambda_f| = 1$, or $\left|\frac{q}{p}\right| = 1$. The second term describes the interference between the decay with the mixing and the decay without the mixing. It vanishes when $\Im m(\lambda_f) = 0$.

1.5 Theoretical methods to measure the γ angle

A good knowledge of the γ angle allows both to test the Standard Model and to probe New Physics (NP) scenarios. As an example, NP events would not obey to the unitarity condition $\alpha + \beta + \gamma = \pi$. Moreover, while α and β have been determined to a good level of precision [21], the measurement of γ is still limited by the smallness of the branching ratios involved in the processes used to measure it, and its relative uncertainty varies between 7 and 15%, depending on the method used to combine the experimental results [19, 20, 21].

Several methods can be used to measure γ [22]:

- $B^\pm \to DK^\pm$ [3]. This method has the smallest theoretical uncertainties, of the order of $\sim 1\%$, since the decay is dominated by tree-level amplitudes, with no penguin contributions;

- $B_{u,d} \to \pi\pi, \pi K$. This method has larger theoretical uncertainties; it can be used to set limits on γ, rather than to make precise measurements;

- $B_d \to (D^*)^\pm \pi^\mp$. They are partially reconstructed decays; they can be used to extract $\sin(2\beta + \gamma)$.

- $B_s \to D_s^\pm K^\mp$. It allows to perform a time-dependent analysis.

---

[3]On the whole text we will use D to indicate either D^0 or $\overline{D}^0$ mesons, except in places where it is explicitly written.

We will see in more detail the first type of decays, where γ appears as the relative weak phase between two amplitudes, the favored $b \to c\bar{u}s$ of the $B^- \to D^0 K^-$ (whose amplitude is proportional to the CKM elements $V_{cb}V_{us}$) and the color-suppressed $b \to u\bar{c}s$ of the $B^- \to \overline{D}^0 K^-$ (whose amplitude is proportional to $V_{ub}V_{cs}$). The interference between D^0 and $\overline{D}^0$ decaying into the same final state leads to measurable CP-violating effects. According to the final state of the D, the following methods have been suggested to infer γ:

- *GLW (Gronau-London-Wyler) method* [23, 24], which uses CP eigenstates of D^0, as $D^0_{CP+} \to K^+K^-, \pi^+\pi^-$ and $D^0_{CP-} \to K^0_s\pi^0, K^0_s\phi, K^0_s\omega$;

- *ADS (Atwood-Dunietz-Soni) method* [1, 2], which uses the doubly Cabibbo suppressed mode $D^0 \to K^+\pi^-$;

- GGSZ (or Dalitz) method [25, 2], which uses three body decays of D^0, as $D^0 \to K^0_s\pi^+\pi^-$.

1.5.1 Measurement of the γ angle through $B^\pm \to DK^\pm$ decays

In Fig. 1.4 is shown the Feynman diagram of the two interfering amplitudes, giving rise to the γ angle. Being the amplitude of the *color suppressed* decay of one order
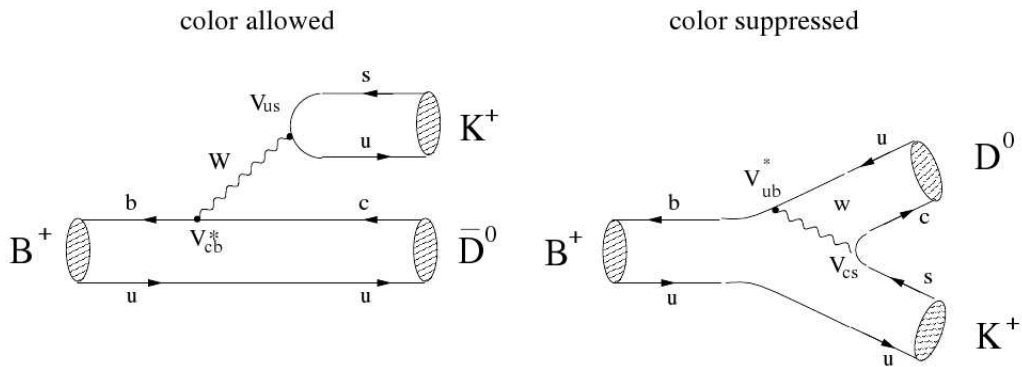
Figure 1.4: Feynman diagram of the decays $B^+ \to \overline{D}^0 K^+$ and $B^+ \to D^0 K^+$.

of magnitude smaller than the *color allowed*, the expected interfering effects are of the order of 10%.

The GLW method

The GLW method measures the *CP*-violating asymmetry in decays of the type $B^+ \to D_{CP_\pm} K^+$ where $CP_\pm$ means that the D decays into a *CP* eigenstate final state.

$$|D_{CP_\pm}\rangle = \frac{1}{\sqrt{2}} \left(|D^0\rangle \pm |\overline{D}^0\rangle \right) .$$

The decay amplitudes can be easily written as:

$$\sqrt{2} A(B^+ \to D_{CP_+} K^+) = A(B^+ \to D^0 K^+) + A(B^+ \to \overline{D}^0 K^+),$$

$$\sqrt{2} A(B^- \to D_{CP_+} K^-) = A(B^- \to D^0 K^-) + A(B^- \to \overline{D}^0 K^-) .$$

Those relations are exact and can be represented as triangles in the complex plane, as shown in Fig. 1.5.
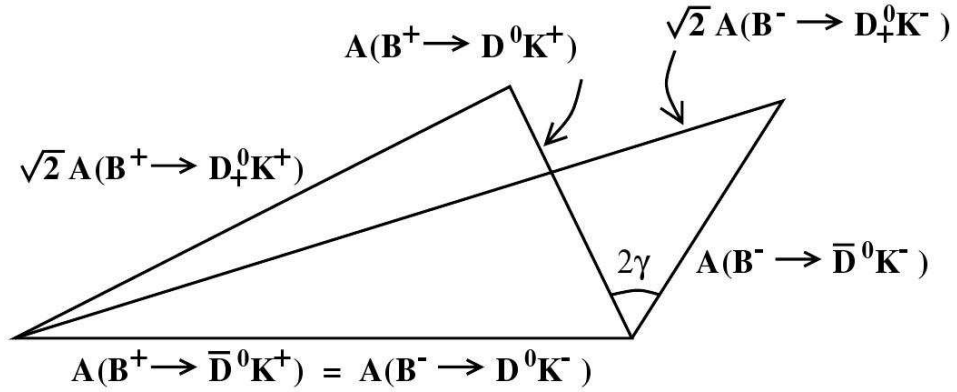
Figure 1.5: Representation in the complex plane of the relations giving the amplitudes of $B^+ \to D_{CP_\pm} K^+$ according to the GLW method.

They can be written also in terms of the CKM matrix elements, considering the quark transitions between $b \to c\overline{u}s$ and $b \to u\overline{c}s$:

$$A(B^+ \to D^0 K^+) = e^{i\gamma} \lambda |V_{cb}| R_B |\mathcal{M}_a| e^{i\delta_a}$$

where γ is the weak phase, $\lambda = V_{cd}$, $R_B = \left| \dfrac{V_{cs} V_{ub}}{V_{cd} V_{cb}} \right|$ is one of the sides of the unitarity triangle (see Fig. 1.2), $\mathcal{M}_a$ is the matrix element of the decay, and δ_a is the strong phase.

1. CP violation and methods for the measurement of the γ angle

If we consider the conjugate process we obtain:

$$A(B^- \to \overline{D}^0 K^-) = e^{-i\gamma} \lambda |V_{cb}| \left| \frac{V_{cs} V_{ub}}{V_{cd} V_{cb}} \right| |\mathcal{M}_a| e^{i\delta_a} \ .$$

From these relations it follows that:

$$A(B^+ \to D^0 K^+) = e^{2i\gamma} A(B^- \to \overline{D}^0 K^-) \ .$$

For the *color allowed* decays we have:

$$A(B^+ \to \overline{D}^0 K^+) = \lambda |V_{cb}| |\mathcal{M}_A| e^{i\delta_A}.$$

where $\lambda = V_{us}$, $\mathcal{M}_A$ is the matrix element of the decay and δ_A is the strong phase. Conventionally we fix this vector along the x axis, so do not introduce γ in its definition. For this reason the amplitude of the conjugate process is the same

$$A(B^+ \to \overline{D}^0 K^+) = A(B^- \to D^0 K^-) \ .$$

If $\gamma \neq 0$

$$\left| A(B^+ \to D^0_{CP_+} K^+) \right| \neq \left| A(B^- \to D^0_{CP_+} K^-) \right|$$

and its value can be clearly extracted measuring the decay rate of the six processes.

One of the disadvantage of this method is that one of the triangle is degenerate, since the amplitude of $B^- \to \overline{D}^0 K^-$ is suppressed with respect to $B^- \to D^0 K^-$ by a factor [26]:

$$r_B \equiv \left| \frac{A(B^- \to \overline{D}^0 K^-)}{A(B^- \to D^0 K^-)} \right| \approx \left| \frac{V_{ub} V_{cs}^*}{V_{cb} V_{us}^*} \right| \left| \frac{a_2}{a_1} \right| \approx 0.1$$

where $\frac{a_2}{a_1} \sim \frac{1}{3}$ indicates the color factor suppression between the *color allowed* and *color suppressed* processes. The experimental value of r_B is: $r_B = 0.107^{+0.010}_{-0.010}$ [19].

The observables for the GLW method are:

$$R_{CP_\pm} \equiv \frac{BR(B^- \to D^0_{CP_\pm} K^-) + BR(B^+ \to D^0_{CP_\pm} K^+)}{BR(B^- \to D^0 K^-) + BR(B^+ \to \overline{D}^0 K^+)} = 1 + r_B^2 \pm r_B \cos \delta_B \cos \gamma,$$

$$A_{CP_\pm} \equiv \frac{BR(B^- \to D^0_{CP_\pm} K^-) - BR(B^+ \to D^0_{CP_\pm} K^+)}{BR(B^- \to D^0_{CP_\pm} K^-) + BR(B^+ \to D^0_{CP_\pm} K^+)} = \frac{\pm 2 r_B \sin \delta_B \sin \gamma}{R_{CP_\pm}},$$

where $\delta_B = \delta_a - \delta_A$ is the difference between the strong phases.

The asymmetry is expected to be small since the interfering amplitudes are of different order of magnitude. A solution to this comes from the ADS method, where decays with comparable amplitudes are used.

The ADS method

The ADS method is a generalization of the GLW method, because it uses decays of the D that are not CP eigenstate. When the decay amplitudes are about of the same order of magnitude, they can give rise to a larger CP-asymmetry, allowing for a better extraction of the γ angle.

Using the notation introduced in [1, 2], we can define the following branching ratios:

$$
\begin{aligned}
a &= \mathcal{B}(B^- \to D^0 K^-) & \overline{a} &= \mathcal{B}(B^+ \to \overline{D}^0 K^+) \\
b &= \mathcal{B}(B^- \to \overline{D}^0 K^-) & \overline{b} &= \mathcal{B}(B^+ \to D^0 K^+) \\
c(f) &= \mathcal{B}(D^0 \to f) & \overline{c}(\overline{f}) &= \mathcal{B}(\overline{D}^0 \to \overline{f}) \\
c(\overline{f}) &= \mathcal{B}(D^0 \to \overline{f}) & \overline{c}(f) &= \mathcal{B}(\overline{D}^0 \to f) \\
d(f) &= \mathcal{B}(B^- \to [f]K^-) & \overline{d}(\overline{f}) &= \mathcal{B}(B^+ \to [\overline{f}]K^+)
\end{aligned}
$$

where f is the common final state between D^0 and $\overline{D}^0$. The interesting case, on which this thesis is focused, is when $f = K^+ \pi^-$. In the SM we assume that: $a = \overline{a}$, $b = \overline{b}$, $c(\overline{f}) = \overline{c}(f)$ and $\overline{c}(\overline{f}) = \overline{c}(\overline{f})$.

A schematic diagram of the two decay channels of interest is reported in Fig. 1.6. Since they are indistinguishable, we are sensitive to their interference, probing γ:

$$
\begin{aligned}
d(f) = \mathcal{B}(B^- \to [f]K^-) &= \mathcal{B}(B^- \to D^0 K^-)\mathcal{B}(D^0 \to f) + \mathcal{B}(B^- \to \overline{D}^0 K^-)\mathcal{B}(\overline{D}^0 \to f) + \\
&\quad + 2\sqrt{\mathcal{B}(B^- \to D^0 K^-)\mathcal{B}(D^0 \to f)\mathcal{B}(B^- \to \overline{D}^0 K^-)\mathcal{B}(\overline{D}^0 \to f)} \cdot \\
&\quad \cos(\delta + \gamma) \\
&= ac(f) + b\overline{c}(f) + 2\sqrt{ac(f) + b\overline{c}(f)}\cos(\delta + \gamma),
\end{aligned}
\tag{1.7}
$$

where $\delta = \delta_B + \delta_D$ is the sum of the strong phases.

Analogous expression is valid also for the conjugate state

$$
\begin{aligned}
\overline{d}(\overline{f}) = \mathcal{B}(B^+ \to [\overline{f}]K^+) &= \mathcal{B}(B^+ \to \overline{D}^0 K^+)\mathcal{B}(\overline{D}^0 \to \overline{f}) + \mathcal{B}(B^+ \to D^0 K^+)\mathcal{B}(D^0 \to \overline{f}) + \\
&\quad + 2\sqrt{\mathcal{B}(B^+ \to \overline{D}^0 K^+)\mathcal{B}(\overline{D}^0 \to \overline{f})\mathcal{B}(B^+ \to D^0 K^+)\mathcal{B}(D^0 \to \overline{f})} \cdot \\
&\quad \cos(\delta - \gamma) \\
&= ac(f) + b\overline{c}(f) + 2\sqrt{ac(f) + b\overline{c}(f)}\cos(\delta - \gamma),
\end{aligned}
\tag{1.8}
$$

which differs from $d(f)$ only for the sign of γ. If CP symmetry is violated, $\gamma \neq 0$, and $d(k, X) \neq \overline{d}(k, \overline{X})$.
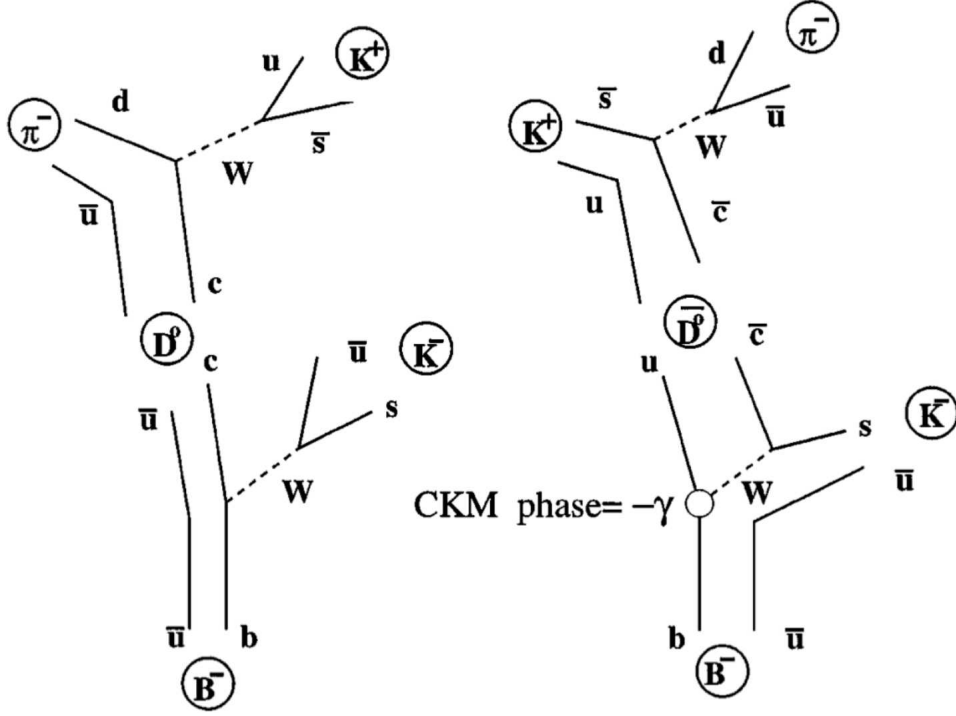
Figure 1.6: Diagram (see [1]) of the two interfering processes: $B^- \to D^0 K^-$ (color allowed) followed by $D^0 \to K^+ \pi^-$ (doubly Cabibbo suppressed) and $B^- \to \overline{D}^0 K^-$ (color suppressed) followed by $\overline{D}^0 \to K^+ \pi^-$ (Cabibbo allowed).

We can compact the square of the sum of the interfering decay amplitudes in this way:

$$d(k, X) = a(k)c(X) + b(k)c(\overline{X}) + 2\sqrt{a(k)b(k)c(X)c(\overline{X})}\cos(\xi_k + \delta_X + \gamma)$$

$$\overline{d}(k, \overline{X}) = a(k)c(X) + b(k)c(\overline{X}) + 2\sqrt{a(k)b(k)c(X)c(\overline{X})}\cos(\xi_k + \delta_X - \gamma)$$

where ξ_k is the difference between the strong phase of $B^- \to K^- D^0$ and $B^- \to K^- \overline{D}^0$, δ_X is the difference of strong phase between $D^0 \to X$ and $D^0 \to \overline{X}$, and γ is the difference of weak phase that violates CP, between $B^- \to K^- D^0$ and $B^- \to K^- \overline{D}^0$.

We define the CP-violating asymmetry

$$A_{ADS}(f) = \frac{d(f) - \overline{d}(\overline{f})}{d(f) + \overline{d}(\overline{f})} = \frac{\mathcal{B}(B^- \to [f]K^-) - \mathcal{B}(B^+ \to [\overline{f}]K^+)}{\mathcal{B}(B^- \to [f]K^-) + \mathcal{B}(B^+ \to [\overline{f}]K^+)} \qquad (1.9)$$

The branching ratio a can be determined via Cabibbo allowed modes of D^0 decay (g), e.g. $D^0 \to K^- \pi^+, K^- \rho^+$. The decay chain $B^- \to D^0[\to g]K^-$ determines $a \approx$

$d(g)/c(g)$ to an accuracy of about 1% since the interfering process $B^- \to \overline{D}^0[\to g]K^-$ is both color and doubly Cabibbo suppressed.

Using $g = K^-\pi^+ = \overline{f}$, we can write

$$a \approx d(g)/c(g) = d(g)/c(\overline{f}),$$

and easily replace $a = d(g)/c(\overline{f})$ and $c(\overline{f}) = d(g)/a$ in (1.7) and (1.8). Calling $r_B^2 = b/a$ and $r_D^2 = c(f)/c(\overline{f})$, we obtain the formulas:

$$d(f) \approx d(g) \cdot [r_B^2 + r_D^2 + 2r_B r_D \cos(\delta + \gamma)] \qquad \boxed{1.10}$$

$$\overline{d}(\overline{f}) \approx \overline{d}(\overline{g}) \cdot [r_B^2 + r_D^2 + 2r_B r_D \cos(\delta - \gamma)] \qquad \boxed{1.11}$$

from which we can define:

$$R^-(K) \equiv \frac{d(f)}{d(g)} = \frac{\mathcal{B}(B^- \to [f]K^-)}{\mathcal{B}(B^- \to [g]K^-)} = \frac{\mathcal{B}(B^- \to [K^+\pi^-]K^-)}{\mathcal{B}(B^- \to [K^-\pi^+]K^-)} \qquad \boxed{1.12}$$

$$R^+(K) \equiv \frac{\overline{d}(\overline{f})}{\overline{d}(\overline{g})} = \frac{\mathcal{B}(B^+ \to [\overline{f}]K^+)}{\mathcal{B}(B^+ \to [\overline{g}]K^+)} = \frac{\mathcal{B}(B^+ \to [K^-\pi^+]K^+)}{\mathcal{B}(B^+ \to [K^+\pi^-]K^+)}. \qquad \boxed{1.13}$$

They are the ratios between branching ratios ($\mathcal{B}$) of suppressed over favored decay modes, for negative and positive charges respectively.

In the same way we can define also:

$$R_{ADS}(K) \equiv \frac{d(f) + \overline{d}(\overline{f})}{d(g) + \overline{d}(\overline{g})} = \frac{\mathcal{B}(B^- \to [K^+\pi^-]K^-) + \mathcal{B}(B^+ \to [K^-\pi^+]K^+)}{\mathcal{B}(B^- \to [K^-\pi^+]K^-) + \mathcal{B}(B^+ \to [K^+\pi^-]K^+)}. \qquad \boxed{1.14}$$

$R^\pm$ have been recently [27] introduced because they provide a set of independent observables, rather than A_{ADS} and R_{ADS} that are strongly correlated.

Using the same definitions for $A_{ADS}(K)$, we can find:

$$\begin{aligned}
R^-(K) &= r_B^2 + r_D^2 + 2r_B r_D \cos(\delta + \gamma) \\
R^+(K) &= r_B^2 + r_D^2 + 2r_B r_D \cos(\delta - \gamma) \\
R_{ADS}(K) &= r_B^2 + r_D^2 + 2r_B r_D \cos(\delta)\cos(\gamma) \\
A_{ADS}(K) &= 2r_B r_D \sin(\delta)\sin(\gamma)/\mathcal{R}_{ADS}
\end{aligned} \qquad \boxed{1.15}$$

The large expected asymmetry comes from the ratio of the two interfering decay amplitudes:

$$\left| \frac{A(B^- \to K^- D^0[\to K^+\pi^-])}{A(B^- \to K^- \overline{D}^0[\to K^+\pi^-])} \right| \approx \left| \frac{V_{cb}V_{us}^*}{V_{ub}V_{cs}^*} \right| \left| \frac{a_1}{a_2} \right| \left| \frac{V_{cd}^*V_{us}}{V_{cs}V_{ud}^*} \right| = \left| \frac{1}{r_B} \right| \left| \frac{V_{cd}^*V_{us}}{V_{cs}V_{ud}^*} \right| \approx 2$$

where $\left| \dfrac{V_{cd}^* V_{us}}{V_{cs} V_{ud}^*} \right| \approx \lambda^2$. Being of the same order of magnitude, we expect large inter-ference effects.

Also, consider that if the strong phase is close to zero, than $\gamma \propto r_B^2$, while for larger values, we have $\gamma \propto r_B$. In this method decays with strong phases, for which the interference is enhanced, are preferred.

The disadvantage lies in the smallness of the suppressed decays. A large sample of them is required to have a significant signal and to make a precise measurement. The suppression factor between suppressed and favored decay is about 10^{-2}. This numbers can be easily calculated from the definition of R_{ADS}, using the current values of $r_B(K) = 0.103^{+0.015}_{-0.024}$ [19], $r_D^2 = (3.80 \pm 0.18) \times 10^{-3}$ [16], $\delta_B = 112^{+12}_{-13}$ [19], $\delta_D = 22^{+11+9}_{-12-11}$ [21], and γ [19, 20].

The same formalism is valid also for other final states, as $K^+ \rho^-$, $K^+ \pi^- \pi^0$, and $K^+ \pi^- \pi^+ \pi^-$. Results from these modes can be used together and combined to di-rectly extract γ. In fact another disadvantage of the ADS method is the presence of two observables and four unknowns. To unequivocally extract γ we should consider either other final states, as mentioned, or combine the results with other methods, as the GLW.

Current experimental results from other experiments [21] than CDF are:

Experiment	$A_{ADS}(K)$	$R_{ADS}(K)$
$B\text{A}B\text{AR}$ [28]$(N(B\bar{B}) = 467M)$	$-0.86 \pm 0.47^{+0.12}_{-0.16}$	$0.011 \pm 0.006 \pm 0.002$
$Belle$ [29]$(N(B\bar{B}) = 772M)$	$-0.39^{+0.26+0.04}_{-0.28-0.03}$	$0.0163^{+0.0044+0.0007}_{-0.0041-0.0013}$
$LHCb$ [30]$(\int \mathcal{L}dt = 0.343\ fb^{-1})$	$-0.39 \pm 0.17 \pm 0.02$	$0.0166 \pm 0.0039 \pm 0.0024$

We also report measurements related to the corresponding $D\pi^-$ modes, since measurable, albeit smaller, γ-dependent asymmetries may also be found in these modes [21]. This decay is suppressed by a factor of about $3 \cdot 10^{-3}$ with respect to its favored mode.

We can note that the maximum possible value of the asymmetry (1.9) is $A_{\max} = 2 r_B r_D / (r_B^2 + r_D^2)$, where r_B can generally be $r_B(K)$ or $r_B(\pi)$. Taking into account the CKM structure of the contributing processes, we expect that $r_B(\pi)$ is suppressed by a factor $|V_{cd}V_{us}/V_{ud}V_{cs}| \sim \tan^2 \theta_C$ with respect to $r_B(K)$, where θ_C is the Cabibbo angle, and we assume the same color suppression factor for both DK and $D\pi$ modes.

With the current values [21, 20, 19, 16] of $r_B(K)$, and $r_B(\pi) \sim 0.005$, we expect $A_{\max}(K) \approx 0.90$ and $A_{\max}(\pi) \approx 0.16$.

Experiment	$A_{ADS}(\pi)$	$R_{ADS}(\pi)$
B_{ABAR} [28]($N(B\bar{B}) = 467M$)	$0.03 \pm 0.17 \pm 0.04$	$(3.3 \pm 0.6 \pm 0.4) \cdot 10^{-3}$
$Belle$ [29]($N(B\bar{B}) = 772M$)	$-0.04 \pm 0.11^{+0.02}_{-0.01}$	$(3.28^{+0.38+0.12}_{-0.36-0.18}) \cdot 10^{-3}$
$LHCb$ [30]($\int \mathcal{L}dt = 0.343\ fb^{-1}$)	$0.09 \pm 0.10 \pm 0.01$	$(4.13 \pm 0.41 \pm 0.40) \cdot 10^{-3}$

The GGSZ method

In this method, decays of $B^- \to D^0 K^-$, with the D^0 decaying into a three-body final state, as $D^0 \to K_S\pi^+\pi^-$, are reconstructed. There are several advantages in this method: first of all we expect large interference effects, due to the presence of resonance terms in the three-body decay; the decays are Cabibbo favored, and they have more statistics than the suppressed decays; furthermore charged decays have a larger reconstruction efficiency, and less background, than neutral states.

The method requires to construct of the *Dalitz plot*, the typical two-dimensional histogram used to study the phase space of a three-body final state. We call p_1, p_2, and p_3 the four-vectors of the final particles. We can define the quantities:

$$s_{12} = (p_1 + p_2)^2,$$

$$s_{13} = (p_1 + p_3)^2.$$

The Dalitz plot has the first quantity on the x axis, the second on the y axis. The available phase space is given by the energy-momentum conservation (see Fig. 1.7):

$$s_{12} + s_{13} + s_{23} = M^2 + m_1^2 + m_2^2 + m_3^2 .$$

The third quantity $s_{23} = (p_2 + p_3)^2$ appears constant on the axis at 45°.

In the particular case of the decay of $D^0 \to K_S\pi^-\pi^+$, $p_1 = p_{K_S}$, $p_2 = p_{\pi^-}$, and $p_3 = p_{\pi^+}$. In this way $s_{12} = s_{K_S\pi^-}$, and $s_{13} = s_{K_S\pi^+}$.

The amplitude of this decay can be written as:

$$A_D(s_{12}, s_{13}) = |A_D(s_{12}, s_{13})|e^{i\delta(s_{12}, s_{13})} .$$
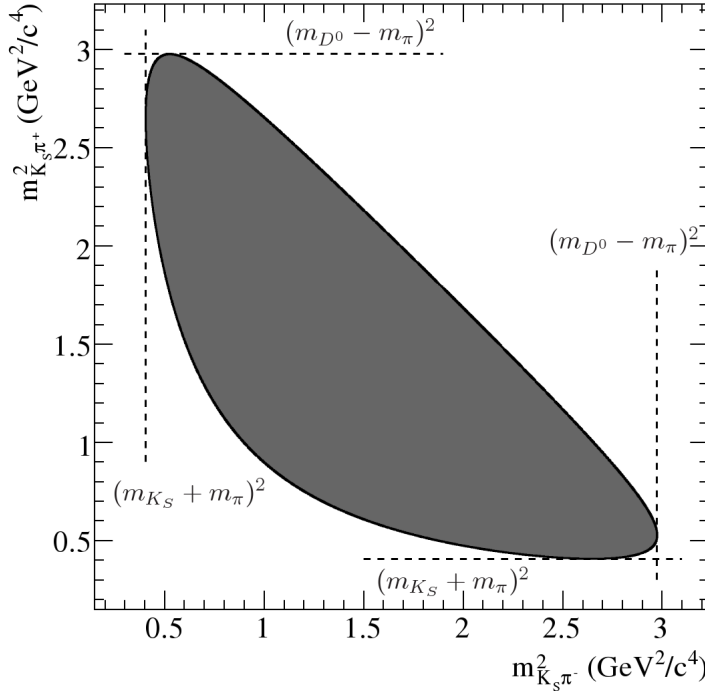
Figure 1.7: Available phase space in the Dalitz plot.

For the B decay we can write:

$$A(B^- \to D^0 K^-) \equiv A_B$$

$$A(B^- \to \overline{D}^0 K^-) \equiv A_B r_B e^{i(\delta_B - \gamma)} \ .$$

The total amplitude of the $B^- \to [K_S \pi^- \pi^-]K^-$ decay, can be written as follows:

$$A(B^- \to [K_S \pi^- \pi^-]K^-) = A_B(A_D(s_{12}, s_{13}) + r_B e^{i(\delta_B - \gamma)} A_D(s_{12}, s_{13})) \ .$$

For the conjugate decay we have:

$$A(B^+ \to [K_S \pi^+ \pi^-]K^+) = A_B(A_D(s_{13}, s_{12}) + r_B e^{i(\delta_B + \gamma)} A_D(s_{13}, s_{12})) \ .$$

In the GGSZ method we need first to construct the Dalitz plot for the D^0 decay, parameterizing its decay amplitude as the sum of some resonant terms and a non-resonant term. In fact it is high the probability that a three-body decay can occur first as a two body decay, where one of the particle, the resonance, immediately decays in other two particles.

$$A_D(s_{12}, s_{13}) = a_{NR}e^{i\phi_{NR}} + \sum_r a_r e^{i\phi_r} \mathcal{A}_r(s_{12}, s_{13}),$$

where r is the index of the resonance and $\mathcal{A}_r(s_{12}, s_{13})$ is the *Breit-Wigner functions*, defined as:

$$\mathcal{A}_r(s_{12}, s_{13}) = {}^J\mathcal{M}_r \times F_{BW}^r,$$

where ${}^J\mathcal{M}_r$ is the term connecting the angular dependence and depends upon the spin J of the resonance. F_{BW}^r is the Breit Wigner relativistic function:

$$F_{BW}^r(s_{ij}) = \frac{1}{s_{ij} - M_r^2 + iM_r\Gamma_r\left(\sqrt{s_{ij}}\right)}$$

where M_r and Γ_r are the mass and width of the r-th resonance.

Fit on the Dalitz plot of the D^0 allows to extract ϕ_i and a_i. To obtain also r_B, δ_B, and γ, we need to create the Dalitz plot also for the $B^\pm \to [K_s\pi^-\pi^+]K^\pm$ decay.

One of the disadvantage of this method comes from the way bins are partitioned in the histogram. They have to be small in order to have a precise measurement, but also large enough to have a significant number of events inside it.

In Fig. 1.8 are reported the combination of the current results and the estimated value of γ, according to different methods. On top there is the result from the CKMFitter Collaboration [19], showing the combination of the $B^- \to DK^-$ methods described above. Also the measurement described in this thesis is already included on it. On the bottom there is the combination proposed by the UTfit Collaboration [20], including other inputs to extract γ. The mean value of γ varies from $67.1^{+4.6}_{-3.7}$ for the first to 74 ± 11 for the latter.
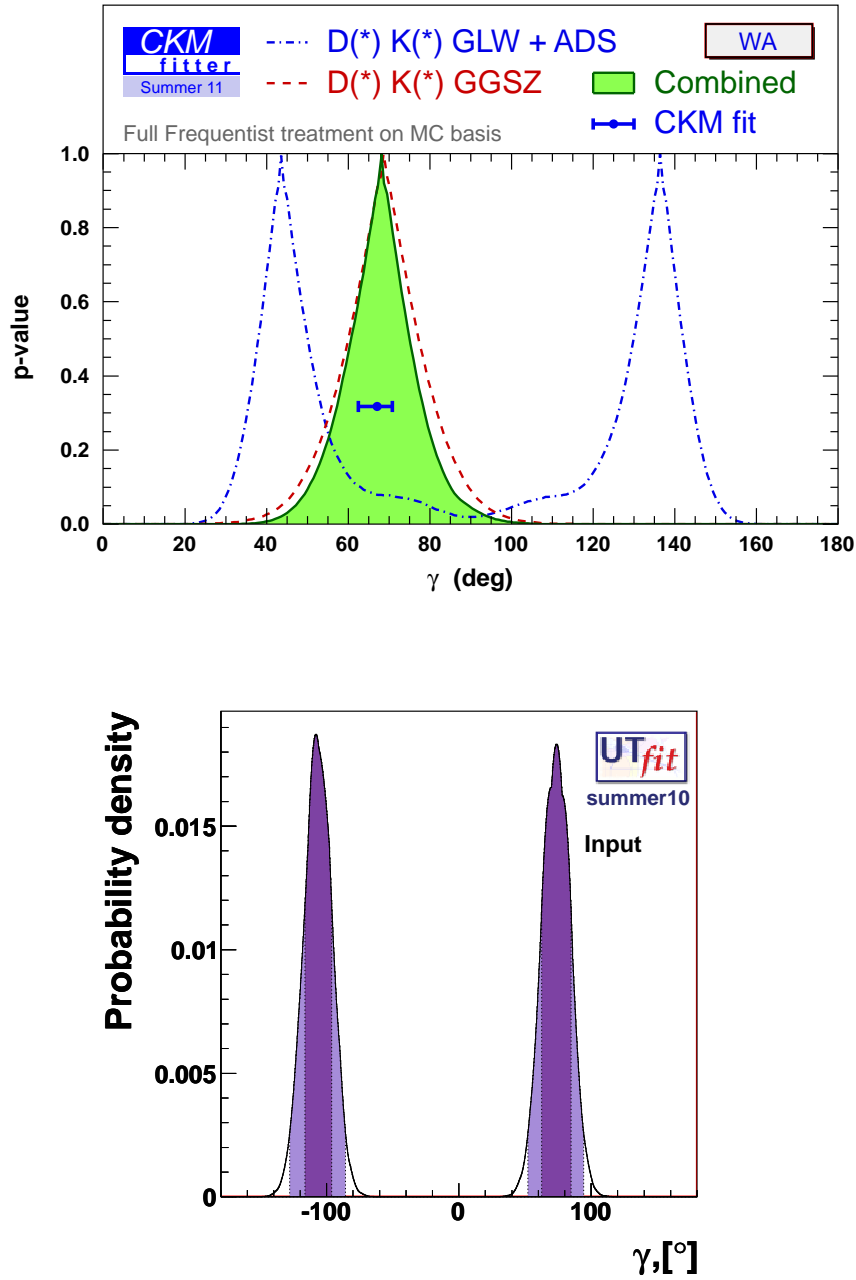
Figure 1.8: Combination of the current results on γ. On top the result from the CKMFitter Collaboration [19], which includes also the result of this thesis. On the bottom the combination proposed by the UTfit Collaboration [20].

2

The Tevatron collider
and the CDF II detector

The measurement described in this Thesis is based on a data sample collected by the CDF II detector during the `Run II` *operations at the Fermilab's Tevatron Collider. In this chapter we will provide a general description of the experimental apparatus, both collider and detector, focusing on the more relevant elements for this analysis.*

2.1 The Tevatron Collider

The last collision at the Tevatron Collider at the Fermi National Accelerator Laboratory (FNAL, or Fermilab) occurred last September 31$^{\text{st}}$ 2011. Before the Large Hadron Collider LHC at CERN have turned on since 2010, Tevatron was the world's highest energy accelerator, colliding protons and anti-protons with a center of mass energy of $\sqrt{s} = 1.96$ TeV.

The first $p\overline{p}$ collision has been produced in 1986, while the accelerator started to collect interesting events at the end of the 1987, during a period called `Run 0`. From then, several extensive upgrades have been undertaken leading to major improvements of the overall performances. `Run I` (Tab. 2.1) went from 1992 to 1996 with a center of mass energy of 1.8 TeV. Then a major upgrade took place between September 1997 and March 2001, for the `Run II` (Tab. 2.1), when both accelerator

and collider detectors were improved.

Date	Event	$\sqrt{s}$ [TeV]	$\mathcal{L}$ [cm^{-2} s^{-1}]	$\int \mathcal{L} dt$ [pb^{-1}]
Mar 1983	End of the construction	-	-	-
Jul 1983	Proton energy: 512 GeV	-	-	-
Oct 1983	Fixed-target program	-	-	-
Feb 1984	Proton energy: 800 GeV	-	-	-
Oct 1985	First $p\overline{p}$ collisions	1.6	10^{24}	-
Oct 1986	Proton energy 900 GeV	-	-	-
Jun 1988–May 1989	Run 0	1.8	2×10^{30}	$\simeq 4.5$
Aug 1992–Feb 1996	Run I	1.8/0.63	28×10^{30}	$\simeq 180$
Aug 2000	Beam energy: 980 GeV	-	-	-
Mar 2001–Sep 2011	Run II	1.96	3.6×10^{32}	$\simeq 10.000$

Table 2.1: Chronological overview of the Tevatron operations and performances.

The upgraded machine accelerates 36 bunches of protons and antiprotons, whereas the previous version of the accelerator operated with only 6. Consequently, the time between bunch crossings has been decreased from 3.5 μs of the previous version to 396 ns for the current collider. The center of mass energy was also increased from 1.8 to 1.96 TeV. At this energy the total cross-section using $p\overline{p}$ beams is maximized with respect to pp beams, as for LHC (Fig. 2.1).

The Tevatron is the last stage of a complex system of accelerators (see Fig. 2.2), used in successive steps to produce, store and accelerate the particles. It is a circular synchrotron of 1 Km radius, employing 772 dipole, 2 halfdipole, and 204 quadrupole superconducting magnets. Each of them is approximately 6 m long, 4 tons in mass, and is made of NbTi alloy filaments embedded in copper, kept at 4.3 K temperature by a large cryogenic system. A 4400 A current flows through each magnet to produce the 4.2 T magnetic field necessary to keep the particles on their orbit, while they are accelerated by eight Radio-frequency cavities (RF) driven at approximately 53.105 Hz.

There are two colliding points, located in two regions along the ring: BØ, which is the site of the CDF experiment, and DØ, where the DØ experiment is located.

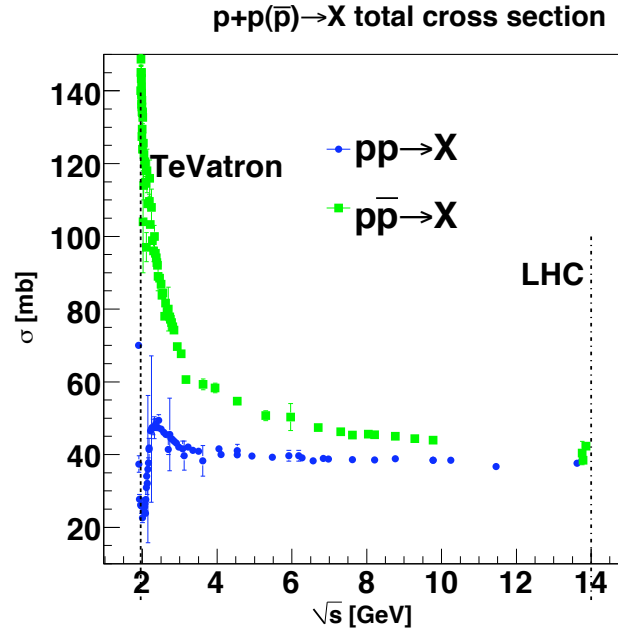**p+p(p̄)→X total cross section**

Figure 2.1: Comparison of the total cross-section as function of the center of mass energy $\sqrt{s}$ for $pp \to X$ (LHC) and $p\bar{p} \to X$ (Tevatron) [16].
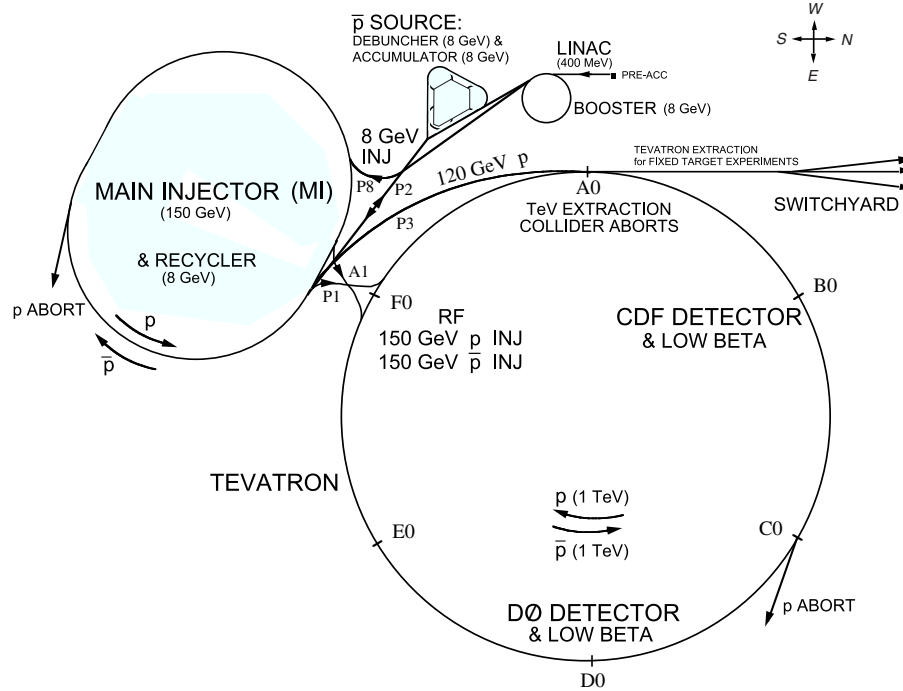

Figure 2.2: The accelerator system operating at FNAL.

2.1.1 Proton production

Protons are extracted from an hot hydrogen gas in the molecular state H_2, ionized by the passage through a magnetron, and accelerated to 750 KeV by a three-staged diode-capacitor voltage multiplier Cockroft-Walton accelerator [31]. The proton beam, segmented into bunches, is then injected into the two-staged 150 m long linear accelerator Linac (see Fig. 2.2). First, a drift tube accelerator accelerates the H^- beams up to 116 MeV; then they are segmented in bunches, of about $5 \cdot 10^{12}$ protons, and the Linac and a Radio Frequency cavity increase their energy to 401.5 MeV [32] before injecting into the Booster. The Booster (see Fig. 2.2) is an alternating gradient synchrotron (orbit radius of 75.5 m) that accelerates protons to 8 GeV. At injection, a thin carbon foil is used to strip the electrons from the H^{-1} ions to obtain protons. Injecting H^{-1} ions rather than protons into the Booster allows the injection to proceed over multiple revolutions of the beam around the Booster Ring (usually $10 \div 12$)[1].

The bunches are injected in the *Main Injector* [33], a synchrotron that accelerates them up to 150 GeV of energy. From the Main Injector the protons move to the Tevatron, a circular accelerator with superconducting magnets of 4.2 T, that accelerates protons up to the final energy of 980 TeV.

2.1.2 Anti-proton production

Anti-proton beams are produced from proton bunches. When the latters reach the energy of 120 GeV they are extracted from the Main Injector and directed against a target made of nickel alloys containing chromium, iron and other metals. The target rotates so that its resistance to the radiation is increased. Anti-protons are produced by the following reaction:

$$p + \begin{pmatrix} n \\ p \end{pmatrix} \quad \rightarrow \quad p + p + \begin{pmatrix} n \\ p \end{pmatrix} + \overline{p} \qquad \boxed{2.1}$$

A spatially broad beam of particles is produced and then focused using a cylindri-

---

[1]If protons were instead injected, the magnetic field used to inject new protons onto orbit in the Booster would also deflect the already revolving protons out of orbit.

cal Lithium Lens[2]. This beam, which has a bunch structure similar to the incident proton beam, is passed through a pulsed dipole magnet. The magnetic field selects the negatively charged antiprotons with about 8 GeV of kinetic energy. About 20 antiprotons are produced for every 10^6 protons on target and then stored into the "Debuncher", a triangular-shaped synchrotron with mean radius of 90 m. The beam is stochastically cooled [34] [3] and then transferred to the "Accumulator", another triangular-shaped synchrotron with a mean radius of 75 m. The Accumulator is a storage ring for the anti-protons; they are stored at an energy of 8 GeV and cooled until needed. The anti-protons are then sent into the Main Injector, where they are accelerated to 150 GeV. Finally, the anti-protons are transferred to the Tevatron, where 36 previously injected bunches of protons are already circulating in opposite direction. Since 2004, an additional Recycler Ring has been added in the same tunnel of the Main Injector and provides additional storage and cooling of the anti-protons. Recently, relativistic electron cooling was successfully implemented in the Recycler, further enhancing the Tevatron performance [35][4]. The anti-proton production rate is low mainly because the production efficiency is $20 \cdot 10^{-6}$, followed by other inefficiency in the transfers.

2.1.3 Injection and collision

The anti-proton accumulation process usually takes about 20 hours to collect a sufficient number of anti-protons, then the accumulation is stopped and the accelerator is prepared for a new injection. The first injection step is the extraction of a set of seven proton bunches from the Booster into the Main Injector [33], where they are accelerated up to 150 GeV. Within the Booster the protons are compacted in a single bunch of $\approx 300 \cdot 10^9$ particles, with an efficiency of 90%. When the proton bunches are ready, they are moved into the Tevatron. This process is repeated every 12.5

---

[2]Lithium is used to minimize beam loss from multiple-scattering.

[3]Stochastic cooling is a technique used to reduce the beam emittance, without any accompanying beam-loss. When the particle deviates from its ideal orbit, an iterative feedback mechanism produces an electrical signal that brings it back.

[4]In the electron cooling process an electron beam propagates at the same average velocity of the particle beam and through Coulomb scattering the anti-protons loose energy until a thermal equilibrium is reached.

s, until 36 proton bunches, separated by 396 ns, are loaded into the Tevatron central orbit, with an efficiency of 65%. The proton injection precedes the anti-proton injection, because if proton bunch losses are large, the proton injection is aborted and restarted. At this stage the process can be quickly recovered, while a large anti-proton lost needs a new accumulation. When the proton bunches are injected and stable, the anti-proton bunches are extracted from the Accumulator (or from the Recycler) to the Main Injector, accelerated to 150 GeV, then compacted into four bunches with an efficiency of $\approx 80\%$. Finally each bunch has $80 \cdot 10^9$ anti-protons.

The four bunches, separated by 396 ns, are injected into the Tevatron, where protons are counter-routing. The two beams have orbits spatially separated of 3-5 mm, that corresponds to $3 - 5\sigma$ of the beam size, to avoid beam interactions outside the collision points. The anti-proton process is repeated to have 36 anti-proton bunches. After the end of anti-protons injection start a "store", defined as an interrupted period of collisions. A store lasts usually about 20-24 hours. The proton and anti-proton bunches share the same magnets and Radio Frequency system. After the injection the beams are accelerated up to 980 GeV in about one minute.

The beams are finally brought into collision at the two instrumented interaction-points located along two straight sections of the Tevatron: DØ and BØ where the DØ and CDF II detectors are located.

Special high-power quadrupole magnets ("low-β squeezers"), installed on the beam pipe at either side of the detectors, reduce the transverse spatial spread of the beams to maximize the collision rate in the interaction regions. The resulting transverse spatial distribution of the luminous region is approximately a two-dimensional Gaussian, with $\sigma_T \approx 30$ μm. The typical longitudinal dimension of a bunch is 60÷70 cm. The interaction regions have a roughly Gaussian distribution along the beam direction, with r.m.s. width $\sigma_z \approx 28$ cm. The center of the luminous region is shifted toward the nominal interaction point by fine tuning of the squeezers. The 36 bunches of protons (antiprotons) are distributed among the 1113 buckets in three equispaced "trains" of 12 bunches each (Fig. 2.3). The inter-bunch spacing is 396 ns (21 buckets) within a train, while a 2.6 μs spacing (139 buckets, called "abort

gap") is kept between trains. The need for the abort gap is two-fold: it allows an-
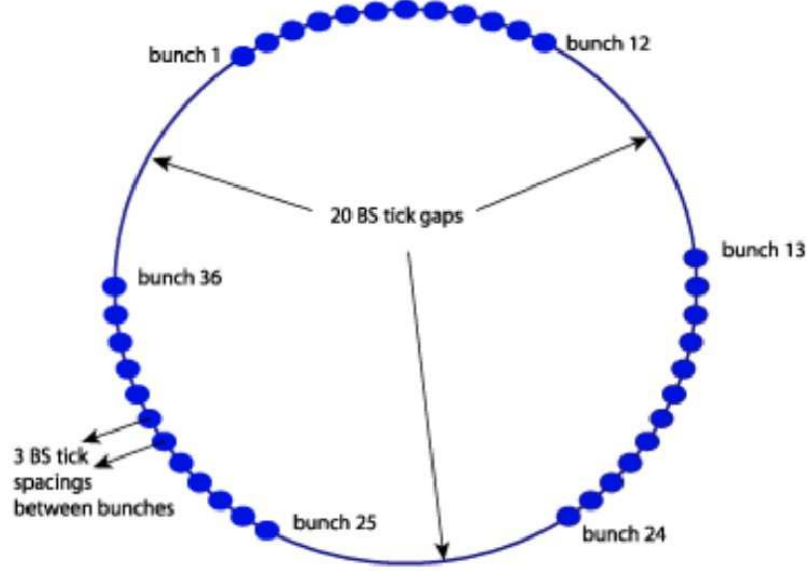

Figure 2.3: Bunch structure of the Tevatron (BS = beam sync ticks = 132 ns).

tiprotons injection (in coincidence with the proton abort gap) without perturbing the already revolving protons with the injecting magnet. Furthermore, when beam abortion is needed, the abort gap allows ramping-up the deflecting magnets without interfering with the beam during the transient, possibly damaging the detectors. As a consequence of this bunch distribution, the average bunch-crossing rate is 1.7 MHz, resulting from a 2.53 MHz rate, when the proton and antiproton trains are crossing, and zero rate in correspondence of the abort gaps. The transverse profile of the beam is shaped to its optimized configuration to avoid detector damage from the tails of the p ($\overline{p}$) distributions interacting with the beam pipe: retractable collimators (iron plates) are moved perpendicularly toward the beam and trim-off the residual halo. When the beam profile is narrow enough and the conditions are safely stable, the detector is powered and the data-taking starts. The number of overlapping inelastic interactions N for each bunch crossing is a Poisson-distributed variable that depends on the instantaneous luminosity. The observed distribution of the multiplicity of interaction vertexes yields $\overline{N} \approx 0.2$, 1.0, 2.0, and 6.0 for respectively, $\mathcal{L} \approx 1 \cdot 10^{31}$, $5 \cdot 10^{31}$, $10 \cdot 10^{31}$, and $30 \cdot 10^{31}$ luminosities [36]. The luminosity

decreases as a function of time during the store because of the interactions of the beam with residual molecules of gas in the beam pipe, beam-halo interactions, and $\overline{p}$ depletion due to the collisions. During the $10 \div 20$ h of a store, the luminosity decreases by a factor of $2.5 \div 5$, the majority of data being collected at $\mathcal{L} \approx \mathcal{L}_0/2$. Just after the final injection, a new anti-proton accumulation cycle starts. When the antiproton stack is sufficiently large and the colliding beams are degraded, the detector high-voltages are switched-off and the store is dumped. The beam is extracted via a switch-yard and sent to an absorption zone. Beam abortion can occur also accidentally when a superconducting magnet rises its temperature above the critical value (i. e., the magnet "quenches"), destroying the orbit of the beams.

The time between the end of a store and the beginning of collisions of the next one is typically 2 hours, during which the calibrations of the sub-detectors and cosmic rays tests are usually performed.

2.1.4 Luminosity

The performances of the Tevatron collider are evaluated in terms of two key parameters: the available center-of-mass energy $\sqrt{s}$, and the instantaneous luminosity $\mathcal{L}$. The former defines the accessible phase-space for the production of resonances in the final states. The latter is defined as the interaction rate per unit cross section of the colliding beams (collisions/(cm^{2}s)) and it is the proportionality coefficient between the rate of a process $\left(\dfrac{dN}{dt}\right)$ and its cross section (σ).

$$\frac{dN}{dt} \; \left[\text{eventi s}^{-1}\right] = \mathcal{L} \; \left[\text{cm}^{-2}\text{s}^{-1}\right] \times \sigma \; \left[\text{cm}^2\right]$$

The time integral of the luminosity is therefore the measurement of the expected number of events N produced in a time T. It is important because it relates both to the peak performances of the accelerator and to the duty cycle of the machine.

Assuming an ideal frontal collision $p\overline{p}$, the instantaneous luminosity is defined as:

$$\mathcal{L} = 10^{-5} \frac{N_p N_{\overline{p}} B f \beta \gamma}{2\pi \beta^\star \sqrt{(\epsilon_p + \epsilon_{\overline{p}})_x (\epsilon_p + \epsilon_{\overline{p}})_y}} \; F\left(\sigma_z/\beta^\star\right) \qquad \left[10^{30}\text{cm}^{-2}\text{s}^{-1}\right] \qquad \boxed{2.2}$$

where N_p ($N_{\overline{p}}$) is the mean number of protons (anti-protons) in each bunch, B (36) is the number of bunches circulating in the ring, f (47.713 kHz) is the revolution

frequency, $\beta\gamma$ is the relativistic factor of the Lorentz boost ($\beta\gamma = 1045.8$ at $\sqrt{s} = 980$ GeV). F is the form factor [5], which depends on the ratio between the longitudinal length of the bunch σ_z and the betatron function [6] $\beta^\star$ evaluated in the intersection point of the bunches ($\beta^\star \approx 31$ cm). ϵ_p ($\epsilon_{\overline{p}}$) is the emittance [7] of the proton (anti-proton) beam, after the injection $\epsilon_p \approx 18\pi$ mm mrad and $\epsilon_{\overline{p}} \approx 13\pi$ mm mrad.

In Table 2.2 there are summarized the characteristic values for the accelerator.

Parameter	Run II value
number of bunches (N_b)	36
revolution frequency [MHz] (f_{bc})	1.7
bunch rms [m] σ_l	0.37
bunch spacing [ns]	396
protons/bunch (N_p)	2.7×10^{11}
antiprotons/bunch ($N_{\bar{p}}$)	3.0×10^{10}
total antiprotons	1.1×10^{12}
β^* [cm]	35

Table 2.2: *Accelerator parameters for Run II configuration.*

Figs. 2.4(a) and 2.4(b) show, respectively, the evolution in the integrated luminosity and the instantaneous luminosity at the start of store delivered by Tevatron. The progressive increase in the integrated luminosity and the continuous records in the instantaneous luminosity [8] proved the good performance of the accelerator.

2.2 The CDF II Detector

The upgraded CDF detector [37] is a large multi-purpose solenoidal magnetic spectrometer designed with an approximately cylindrically symmetric layout both in

---

[5]The form factor F describes the longitudinal profile of the beam in the collision region, assuming the characteristic shape of an horizontal hourglass, centered in the interacting point.

[6]The betatron function is a measure of the beam width, along the accelerator and it is proportional to the beam's x and y extent in phase space. β^* is the value of this function in the collision point
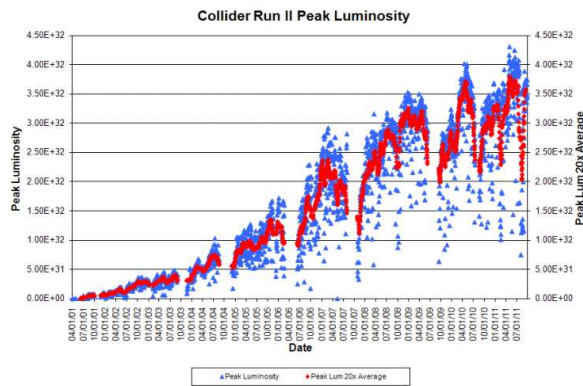
[7]The emittance ϵ measures the phase space occupied by the beam particles

[8]In February 2010, the record in the instantaneous luminosity was $4.0 \times 10^{32} cm^{-2} s^{-1}$.
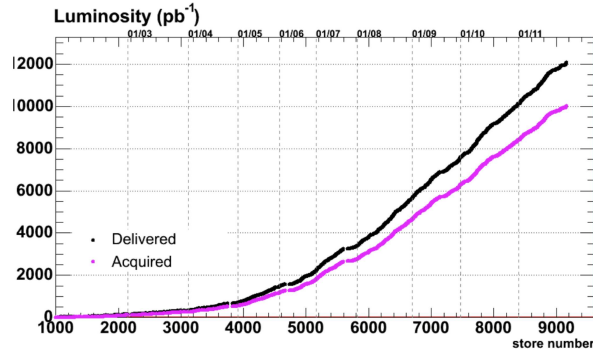
(a)

(b)

Figure 2.4:   **2.4(a):** Integrated luminosity as a function of the time (or store number). The black curve is the luminosity delivered and the purple curve is luminosity written to tape. **2.4(b):** Initial luminosity as a function of the time (or store number).

the azimuthal plane and in the "forward" and "backward" [38, 39] directions to study collisions at Tevatron. Fig. 2.5 shows the CDFII detector and the different sub-systems in a solid cut-away view, while Fig. 2.6 shows the elevation view of one half of the detector.

CDF II consists of five main subdetector system: tracking, particle identification, calorimetry, muon identification and luminosity detector. The protons and anti-protons beams travel towards each other along the horizontal axis (beam line or beam axis). Any plane perpendicular to the beam line is called a *transverse plane* and the intersection point between the beam line and the transverse plane is referred to as a *beam spot*.

The innermost system is the integrated tracking system: a silicon microstrips
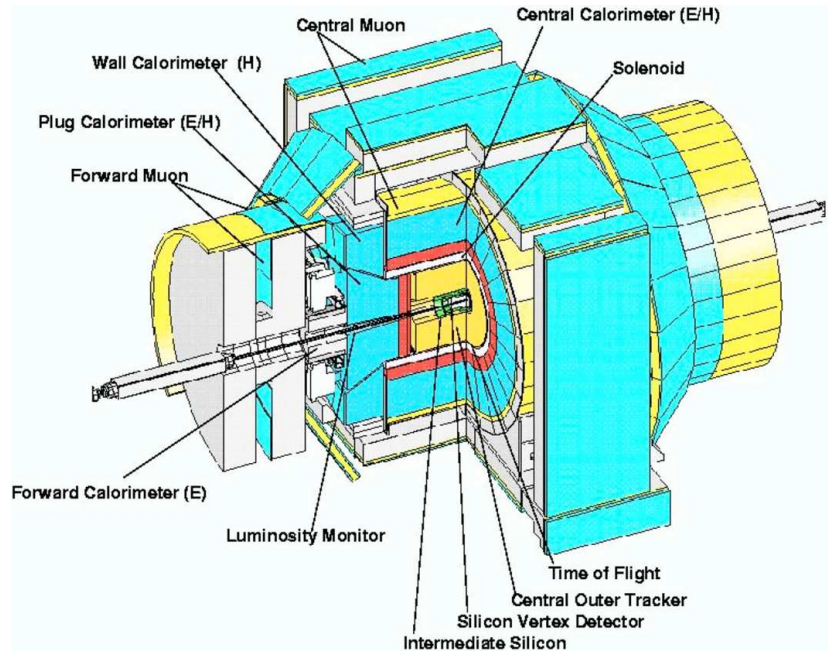
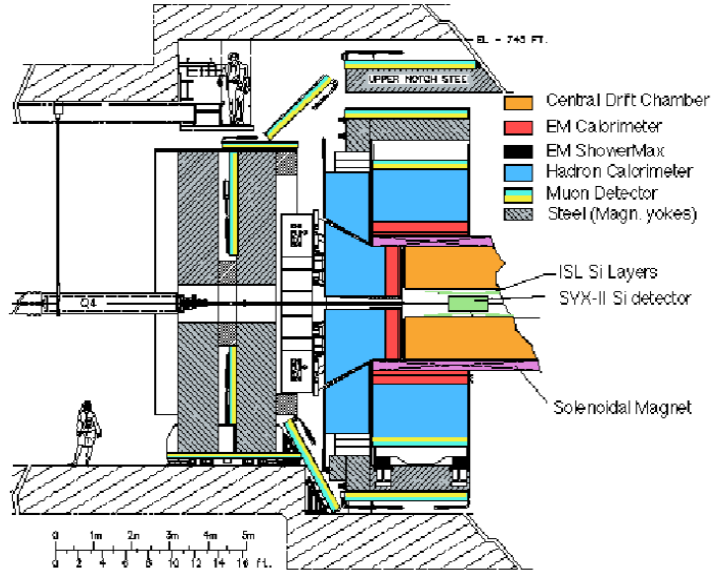Figure 2.5: The CDFII detector with a quadrant cut to expose the different subdetectors.

Figure 2.6: Elevation view of one half of the CDFII detector.

detector and an open-cell wire drift chamber, the Central Outer Tracker COT that surrounds the silicon detector. The tracking system is surrounded by the Time Of Flight detector TOF, designed to provide particle identification for low-momentum charged particles, below 2 GeV/c.

Both the tracking and the TOF detector are placed inside the supeconducting coil, which generates a 1.4 T uniform horizontal magnetic field along the z axis inside the tracking volume. The trajectories of the charged particles inside the tracking volume are helixes. The tracking system is designed to measure the momentum and the trajectory of the charged particles. Multiple-track reconstruction allows to identify the vertices where either the $p\overline{p}$ interaction took place (primary vertex) or the decay of a long-lived particle took place (secondary or displaced vertex).

The solenoid coil is surrounded by the calorimeters, which measure the energy of particles that shower when interacting with matter. The calorimeters are surrounded by the muon detectors. Muons are "minimally ionizing particles" that only deposit small amounts of ionization energy in the material. Therefore, they are able to penetrate both the tracking and calorimeter systems. The integrated material of the tracking system, of the TOF detector, of the solenoid and of the calorimeter serves as a particle filter. Particles which penetrate through all that material are mostly muons, and they are detected by the tracks left in the muon chambers, located outside the calorimeter. At the extreme forward region of the CDFII detector two modules of Cherenkov Luminosity Counters (CLC) [40] are placed. They point to the center of the interaction region to record the number of $p\overline{p}$ interactions and determine the instantaneous luminosity.

In the analysis described in this Thesis the tracking system and the trigger are essential components, and they will be described in detail in the following sections. Less details will be given for the remaining systems. More information can be found in [41] and [42].

2.2.1 Coordinates and Standard Definitions in CDF

CDF adopts a left handed cartesian coordinate system with origin at the nominal BØ interaction point, coincident with the center of the drift chamber (see Sec. 2.3.2). The positive z-axis lies along the nominal beam-line and goes in the direction of the proton beam (east). The $x - y$ plane is therefore perpendicular to the beam-line, with the y-axis pointing upward and x-axis in the horizontal plane, pointing radially outward with respect the center of the accelerator ring (Fig. 2.7).

Since the colliding beams of the Tevatron are unpolarized, the resulting physical observations are invariant under rotations around the beam line axis. Also, the approximately cylindrically symmetry of the detector makes the cylindrical coordinate system (r, ϕ, z) particularly convenient to describe the detector geometry.
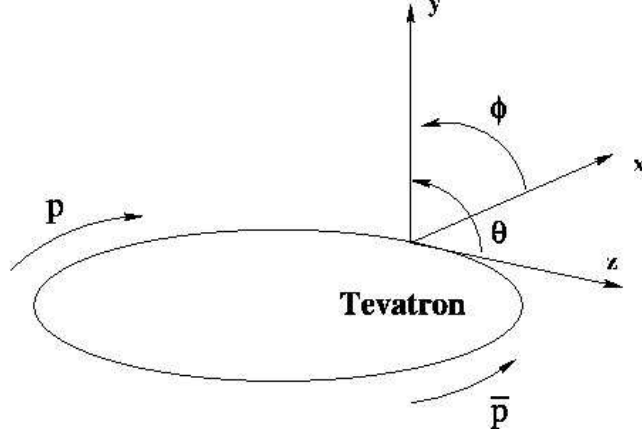

Figure 2.7: CDF coordinate system.

We define the azimuthal angle ϕ, and the polar angle θ:

$$\phi = \tan^{-1} \frac{y}{x}, \qquad \theta = \tan^{-1} \frac{\sqrt{x^2 + y^2}}{z}.$$

A momentum-dependent particle coordinate named *rapidity* is also commonly used:

$$Y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z},$$

where E is the energy and p_z is the z component of the momentum of the particle. Being Lorentz invariant, it is preferred to the polar angle θ. In the relativistic limit, or when the mass of the particle is ignored, rapidity becomes dependent only upon the production angle of the particle with respect to the beam axis. This approximation is called *pseudo-rapidity* η and is defined by

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \tag{2.3}$$

A value of $\theta = 90°$ would be perpendicular to the beam axis and corresponds to $\eta = 0$. The pseudo-rapidity is commonly used to identify different detector regions according to their position with respect to the beamline and interaction vertex position, as shown in Fig. 2.8(a).

Additional quantities are defined, as the *transverse momentum*, $p_T = p \cdot \sin\theta$, the *transverse energy* E_T, and the *five parameter of the helical* trajectories of charged particles moving through a homogeneous solenoidal magnetic field along the z direction. Knowing that the projection of the helix on the $x - y$ plane is a circle, a helix in three dimensions is uniquely parameterized by:

C – signed helix (half)-curvature, defined as $C = q/2R$, where R is the radius of the helix and q is the charge of the track. This is directly related to the transverse momentum $p_T = 0.15 \, qB/|C|$, where the magnetic field B is measured in Tesla, C in m^{-1} and p_T in GeV:

ϕ_0 – ϕ angle of the particle at the point of closest approach to the z-axis;

d_0 – signed impact parameter, i.e. the radial distance of closest approach to the z-axis, defined as $d_0 = q(\sqrt{x_0^2 + y_0^2} - R)$, where (x_0, y_0) are the coordinates of the center. This is schematically drawn in Fig. 2.8(b);

λ – the helix pitch, i.e. $\cot(\theta)$, where θ is the polar angle of the particle at the point of its closest approach to the z-axis. This is directly related to the longitudinal component of the momentum: $p_z = p_T \cot\theta$;

z_0 – the z coordinate of the point of closest approach.

2.3 The Tracking System

The tracking system is immersed in a 1.4 T solenoidal magnetic field for the measurement of charged particles momenta. We will describe this system, shown in Fig. 2.9, starting from the device closest to the beam and moving outwards.

The innermost tracking device is a silicon detector, which consists of three sub-detectors that cover the region $|\eta| < 2$ and 2π of azimuthal angle. The first layer of silicon sensors, called Layer 00 (L00) [43], is installed directly onto the beryllium beam pipe, with the sensors at radii 1.35 and 1.62 cm from the beam [9]. The L00 is

---

[9]The beam pipe is made of beryllium because this metal has the best mechanical qualities with the lowest nuclear interaction cross section.
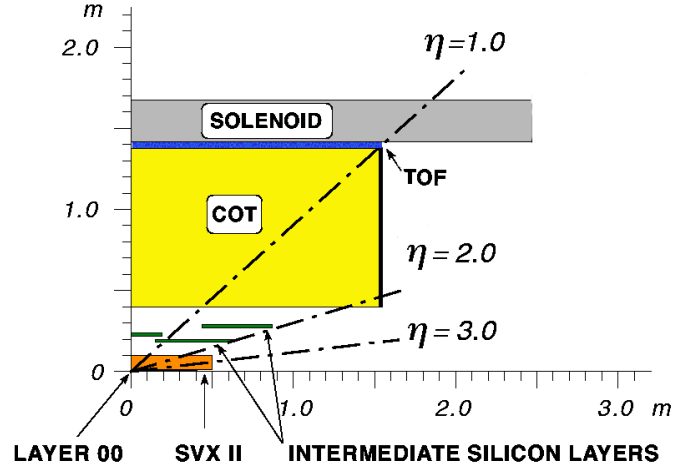
(a)
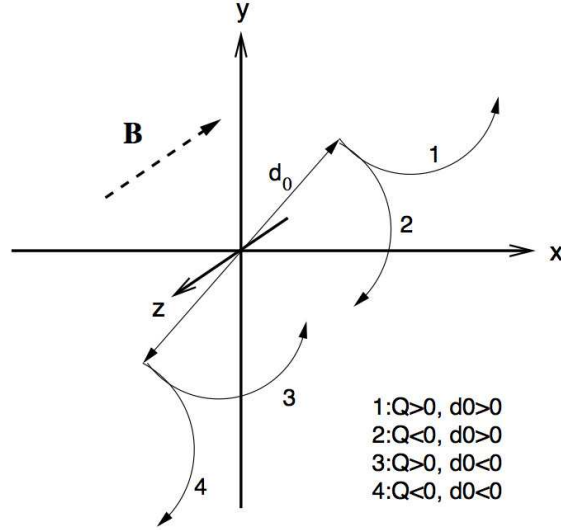
(b)

Figure 2.8: **2.8(a):** The CDFII tracker layout showing the different subdetector systems. **2.8(b):** Schematic drawing of the impact parameter d_0.

followed by the Silicon Vertex Tracker, SVXII [44], made of five concentric layers of silicon sensors located at radii between 2.45 and 10.6 cm. The Intemediate Silicon Layers (ISL) [45] are the outermost silicon detectors, with one layer of sensors at a radius of 22 cm in the central region and two layers at radii 20 and 28 cm in the forward region. Surrounding the silicon detectors is the Central Outer Tracker (COT) [46], a 3.1 m long cylindrical open-cell drift chamber covering the volume between 43.4 cm and 132.3 cm of radius and $|\eta| < 1$.
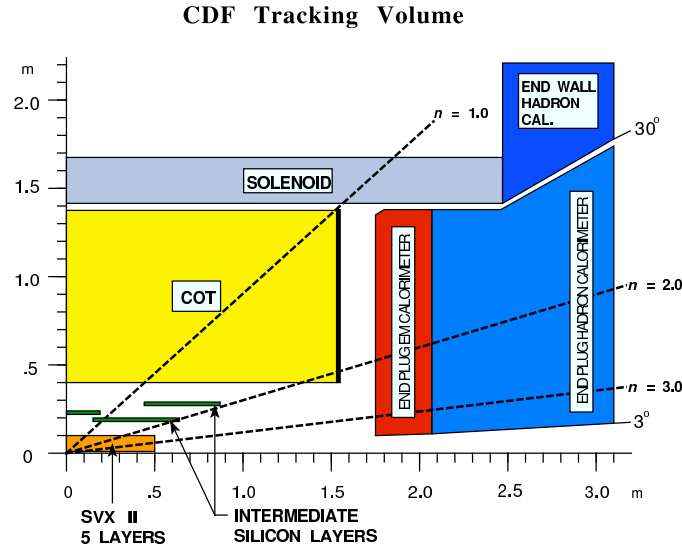
CDF Tracking Volume

Figure 2.9: Elevation view of one quadrant of the inner portion of the CDF II detector showing the tracking volume surrounded by the solenoid and the forward calorimeters.

2.3.1   The Silicon Detectors

The silicon strip detectors [47] provide a precise determination of the particle trajectory close to the beam line. The impact parameter resolution measured in the transverse plane is of 27 μm. A silicon detector is fundamentally a reverse-biased $p - n$ junction.

When a charged particle passes through the detector material, it causes ionization. For a semiconductor, this means that electron-hole pairs are produced. Electrons drift towards the anode, and holes drift towards the cathode, where the charge is gathered. The amount of charge is, to first order, proportional to the path length traversed in the detector material by the charged particle. By segmenting the p or n side of the junction into "strips" and reading out the charge deposition separately on every strip, we obtain sensitivity to the position of the charged particle. All the CDFII silicon detectors are implemented as microstrip detectors. The typical distance between two strips is about 60 μm. Charge deposition from a single particle passing through the silicon sensor can be read out on one or more strips. This charge deposition is used to determine the hit position in the direction perpendicular to the strips.

There are two types of microstrip detectors: single and double-sided. In single-

sided detectors only one p side of the junction is segmented into strips. Double-sided detectors have both sides of the junction segmented into strips. The benefit of double-sided detectors is that while one p side has strips parallel to the z direction, providing $r - \phi$ position measurements, the n side can have strips at an angle (stereo angle) with respect to the z direction, and can provide z position information.

For SVX II, made of double sided silicon sensor, four silicon sensors are assembled into a "ladder" structure which is 29 cm long. The readout electronics are mounted directly to the surface of the silicon sensor at each end of the ladder. The ladders are organized in an approximately cylindrical configuration, creating "barrels". A SVX II barrel is segmented into 12 wedges, each covering approximately $30°$ in ϕ, and for each wedge there are five layers. Each layer provides one axial measurement on one side and a measurement at the stereo angle on the other side (see Tab. 2.3).

Property	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4
number of ϕ strips	256	384	640	768	869
number of z strips	256	576	640	512	869
stereo angle	$90°$	$90°$	$+1.2°$	$90°$	$-1.2°$
ϕ strip pitch [μm]	60	62	60	60	65
z strip pitch [μm]	141	125.5	60	141	65
active width [mm]	15.30	23.75	38.34	46.02	58.18
active length [mm]	72.43	72.43	72.38	72.43	72.43

Table 2.3: Relevant parameters for the layout of the sensors of the five SVXII layers.

The resolution on the single hit is 12 μm. There are three SVXII barrels, mounted adjacent to each other along the z-axis, as shown in Fig. 2.10, covering the nominal interaction region at the center of the CDFII Detector. The coverage of the silicon detector subsystem is shown in Fig. 2.11. The innermost layer, L00, is made of single-sided silicon sensors which provide only $r - \phi$ measurements, but also, being only at 1.5 cm from the interaction point, provides the best resolution on the transverse impact parameter. The ISL is made of double-sided silicon sensors and it provides up to two additional tracking layers depending on pseudorapidity (Fig. 2.11). In particular ISL provides a higher tracking efficiency by connecting tracks in SVX with
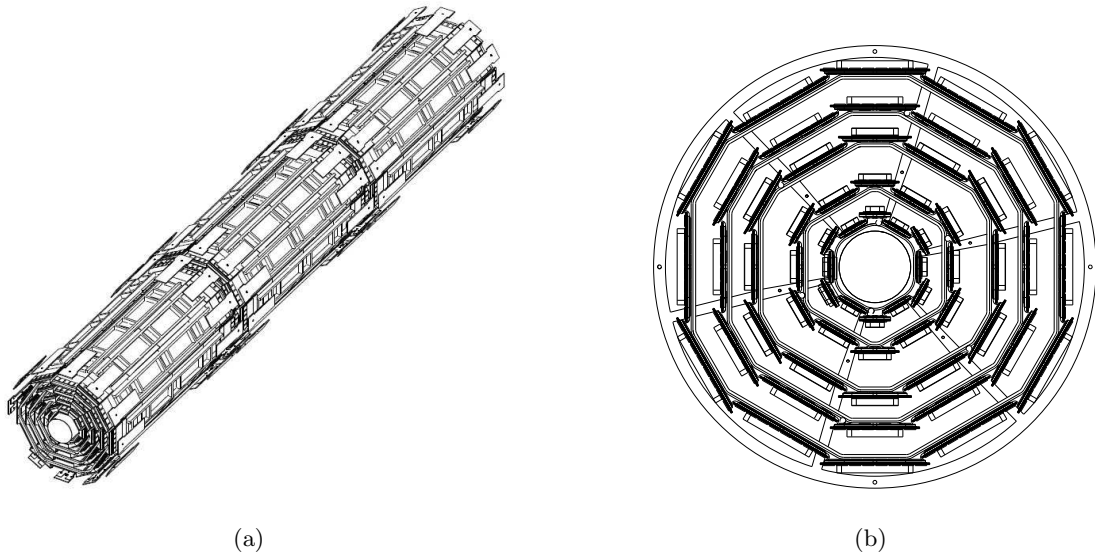
(a)

(b)

Figure 2.10: The SVXII silicon detector. **2.10(a):** Three-dimensional view of the detector, is possible to see the barrel structure along the beam axes. **2.10(b):** The transverse plane section shows in detail the layers sequence.

the ones in COT and allows to extend tracking beyond the COT limit ($|\eta| < 1$), and up to $|\eta| < 2$. All the silicon detectors are used in the Offline track reconstruction algorithm.
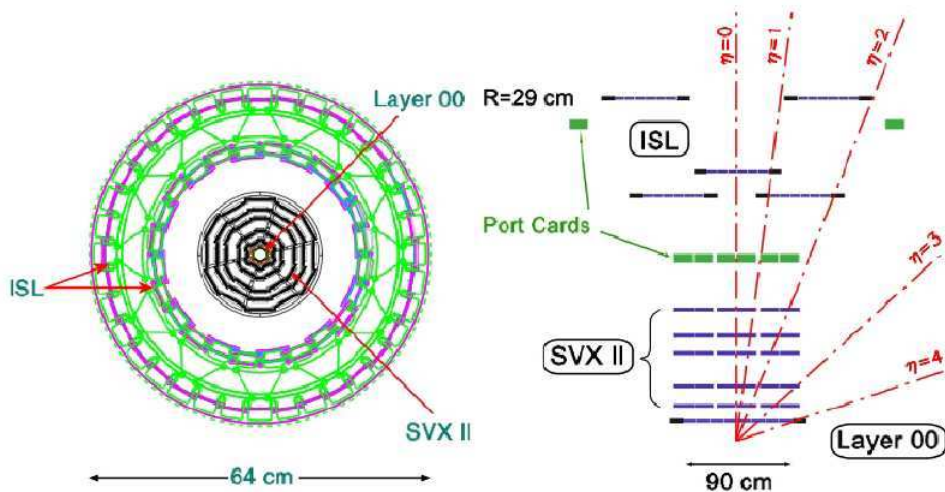

Figure 2.11:   Silicon Detectors: $x - y$ and $z - y$ plane views.

2.3.2 The Central Outer Tracker (COT)

The COT drift chamber provides the tracking of charged particles at radii in the pseudo-rapidity region $|\eta| < 1$, giving an accurate information in the $r - \varphi$ plane for the measurement of the transverse momentum, and substantially less accurate information in the $r - z$ plane for the measurement of the z component p_z. The COT contains 96 sense wire layers, which are radially grouped into eight *superlayers*. This can be seen from the end plate section shown in Fig. 2.12(a). Each superlayer is

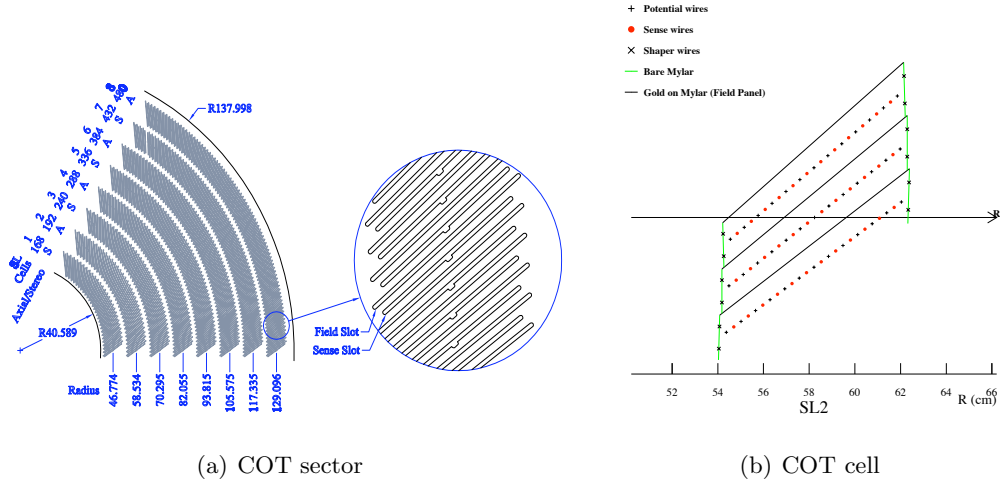|	
(a) COT sector	(b) COT cell

Figure 2.12:   A 1/6 section of the COT end-plate 2.12(a). For each super-layer is given the total number of cells, the wire orientation (axial or stereo), and the average radius in cm. The enlargement in Fig. 2.12 shows in details the slot where the wire planes (sense and field) are installed. The arrow shows the radial direction.

divided into cells, and each cell contains 12 sense wires. The maximum drift distance is approximately the same for all superlayers. Therefore, the number of cells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains $30,240$ sense wires spanning the entire length of the detector in z. Approximately half the wires run along z direction (*axial*). The other half are strung at a small angle ($2°$) with respect to the z direction (*stereo*). This allows to perform track reconstruction in the $r - z$ plane. The active volume of the COT begins at a radius of 43.4 cm from the beamline and extends out to a radius of 132.3 cm. The chamber is 310 cm long. Particles originating from the interaction point with $|\eta| < 1$ pass through all the 8 superlayers of the COT.

The cell layout, shown in Fig. 2.12(b) for superlayer 2, consists of a wire plane containing sense and potential wires (for field shaping) and a field (or cathode) sheet on either side of the cell. Both the sense and potential wires are 40 μm diameter gold plated tungsten wires. The field sheet is 6.35 μm thick mylar with vapor-deposited gold on both sides. Each field sheet is shared with the neighboring cell. The COT is filled with an Argon-Ethane gas mixture and Isopropyl alcohol (with proportions 49.5:49.5:1). The gas mixture is chosen to have a constant drift velocity across the cell width.

When a charged particle passes through the detector volume, the gas is ionized. Electrons drift towards the nearest sense wire. The electric field in a cylindrical system grows exponentially with decreasing radius. As a consequence, an avalanche multiplication of charge happens inside the high electric field region, in the vicinity of the wire, due to *electron-atom* collisions. The resulting charge reaches the wire, producing the so called *hit*, which is then read by electronics. The avalanche discharge provides a gain of ~ 104. The maximum electron drift is approximately 100 ns. Due to the magnetic field, electrons drift at a Lorentz angle of $\sim 35°$ with respect to the radius. The cell is tilted by $\sim 35°$ with respect to the radial direction to compensate for this effect.

Signals on the sense wires are processed by the ASDQ (Amplifier, Shaper, Discriminator with charge encoding) chip, which provides input projection, amplification, pulse shaping, baseline restoration, discrimination, and charge measurement [48]. The pulse is sent through ~ 11 m of micro-coaxial cable, via repeater cards, to TDC boards in the collision hall. The pulse leading edge gives the arrival time information and the pulse width, and it is related to the amount of charge collected by the wire. After calibrating the width variations due to the COT geometry, to the path length of the associated track, and to the gas gain differences for the 96 wires, the Landau function associated to the track is determined, using the amount of charge collected (in nanosecond) for each hit along the track path length. From the Landau distribution the energy loss is measured and used for particle identification. A detailed description of the calibration is found in [49], [50].

The TDC boards contain also the buffer where the data are stored while waiting

for the events to be accepted by the trigger. The TDC auxiliary card catch hits for the XFT track trigger processor (see Sec. 2.9.1), later processed by pattern recognition (tracking) software to form helical tracks. The hit resolution of the COT is about 40 μm. The transverse momentum resolution has been measured using cosmic ray events to be:

$$\frac{\sigma_{p_T}}{p_T^2} = 0.0017 \text{GeV/c}^{-1}. \qquad (2.4)$$

The tracking algorithms reconstruct particle trajectories (helixes) that best correspond to the observed hits. Reconstructed trajectories are referred to as *tracks*. Tracks with available COT information are important for several reasons:

- for the trigger, based on charged tracks and for the special Level-2 (L2) trigger optimized for *B*-physics (see Sec. 2.8.2) used to collect data analysed in this Thesis;

- for the TOF reconstruction to provide particle identification information for the track parent particle;

- for the silicon reconstruction to match the hits in the SVX detector to the COT track trajectory;

- they, themselves, contain information about particle velocity through the measurement of the energy loss.

All the tracks that we use in this analysis are required to have the COT and the SVX II information.

Calibrations of the COT.

The COT was originally designed for an accurate and efficient tracking, but not being explicitly optimized for dE/dx measurements. For example, the small drift cells, designed for fast response and high spatial-resolution measurements, reduce the number of charge clusters collected by each wire, increasing statistical fluctuations. In addition, the gas is kept at atmospheric pressure, while higher pressures are better for dE/dx measurements. However, a dedicated charge-integration mode can be enabled in the COT read-out chip for a measurement of dE/dx .

Off-line dE /dx information could not be used directly in our analysis, because the responses are not uniform both in time and in the chamber volume. In fact a degradation of the sense wires over time has been detected. For this reasons there is the need of a calibration that removes any spurious dE/dx dependence on local properties of the chamber or on features of the particles other than velocity. Moreover greater performances of the particle identification can be achieved after calibrating.

The calibration is made using a sample of pure kaons and pions from the D^0 decay coming from:

$$D^*(2010)^+ \to D^0 \pi^+ \to [K^- \pi^+] \pi^+,$$

where the identity of the decay products of the D^0 is determined by the charge of the pion from D^{*+}.

The steps are first to search for possible dE/dx gain variations as functions of several global and track-related quantities, as the run number, the azimuthal angle, the pseudo-rapidity and the hit multiplicity. Then corrections are applied to remove the dE/dx dependence on these variables.

Finally the ionization curve, the function describing the expected average dE/dx for a charged particle as a function of its Lorentz boost ($\beta\gamma$), is extracted. This curve is particularly useful for particle identification.

In fact, as can be seen in Fig. 2.13, where there is show the dE/dx response as a function of the momentum (Universal Curves), there is a separation among kaons, pions, protons and electrons. For the analysis presented in this Thesis it is important the separation between kaons and pions, that is about 1.5 σ for particles with momentum greater than 2 GeV/c.

2.3.3 Track Reconstruction

As explained in the previous chapter, charged particles leave small charge depositions as they pass through the alternative layers of the tracking system. Using these depositions, pattern recognition algorithms reconstruct the particle original trajectory measuring the five parameters of the helix (see Sec. 2.2.1) that best match to the observed hits.

(a) Universal Curves for positive charges   (b) Universal Curves for positive charges
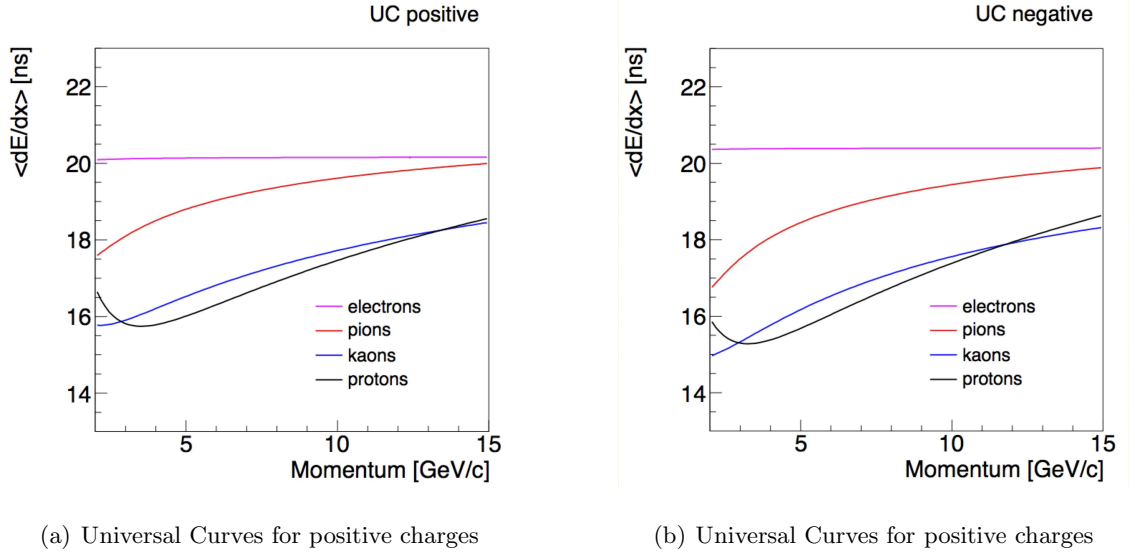
Figure 2.13: Universal Curves for positive 2.13(a) and negative 2.13(b) charges. The distributions for kaons, pions, protons and electrons are overlaid.

CDF employs several algorithms for track reconstruction depending on which component of the detector a particle travels through. The principal one is the Outside-In (OI) reconstruction [51]. This algorithm, exploiting the information from both the central drift chamber and the silicon detectors, is used to track the particles in the central region ($|\eta| < 1$). It first reconstructs tracks in the COT and then extrapolates them inwards toward the beam.

The first step of the pattern recognition in the COT looks for circular paths, projection of helix in the $r - \varphi$ plane, in the axial superlayers. Cells in the axial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called *segments*. Once segments are found, there are two approaches to track finding [52]. One is to link together the segments which are consistent with lying tangent to a common circle. The other approach is to constrain its circular fit to the *beamline* (see Sec. 2.3.4). Once a circular path is found in the $r - \varphi$ plane, segments and hits in the stereo superlayer are added depending on their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. This arises in high track reconstruction efficiency in the COT for tracks passing through all 8

superlayers (97% for tracks with $p_T > 10$ GeV/c)[10].

Once a track is reconstructed in the COT, it is extrapolated inward to the silicon system. Based on the estimated errors on the track parameters, a three dimensional "*road*" is formed around the extrapolated track. Starting from the outermost layer, and working inwards, silicon hits found inside the road are added to the track. As hits get added, the road gets narrowed according to the knowledge of the up-dated track parameters and their covariance matrix. Reducing the width of the road reduces the chance of adding a wrong hit to the track, and also reduces the computation time. In the first step of this algorithm, axial hits are added. In the second step, hits with stereo information are added to the track. At the end, the track combination with the highest number of hits and lowest χ^2/ndf for the five parameters helix fit is kept.

Due to the limited COT coverage and the strict hits requirement (at least four of eight superlayers), tracking in the forward regions requires different algorithms [53], [54] that are not described here because the tracks in the forward regions are not used in this analysis.

2.3.4 Primary Vertex

In this Thesis the location of the primary $p\bar{p}$ vertex is required to calculate the decay length corresponding to the secondary vertices candidates and the impact parameter of charged tracks.

The primary vertex location, for a given event, is found by fitting high quality tracks to a common point of origin. At high luminosities, multiple collisions occur on a given bunch crossing. For a luminosity of 10^{32} cm^{-2}s^{-1}, there are an average of 2.3 interactions per bunch crossing. Typically, since the luminous region is sufficiently long (with $\sigma_z = 29$ cm), the primary vertices associated to the collisions are well separated in z. An iterative algorithm is used to find the vertex associated to the hardest collision: the first estimate of its position (x_V, y_V, z_V) is binned in the

---

[10]The track reconstruction efficiency mostly depends on how many tracks are reconstructed in the event. If there are many tracks close to each other, hits from one track can shadow hits from the other track, resulting in efficiency losses.

z coordinate, then the z position of each vertex is calculated from the weighted average of the z coordinates of all tracks within 1 cm of the first iteration vertex, with a typical resolution of 100 μm; finally the vertex associated with the highest sum of the tracks p_T is defined as primary vertex of the event.

The locus of all primary vertices defines the beamline, the position of the luminous region of the beam-beam collisions through the detector. A linear fit to (x_V, y_V) vs z_V yields the beamline for each stable running period. The beamline is used as a constraint to refine the knowledge of the primary vertex in a given event. Typically the beam transverse section is circular with width of ≈ 30 μm at $z = 0$, rising to $\approx 50 - 60$ μm at $|z| = 40$ cm. The beam is not necessarily parallel nor centered in the detector and moves as a function of time.

2.4 Time Of Flight detector (TOF)

The TOF [55] is a cylindrical array made of 216 scintillating bars and is located between the external surface of the COT and the magnet cryostat containing the superconducting solenoid. It was added in 2001 to improve the capability to distinguish different kind of long-lived particles, measuring the time elapsed between the collision time and when a particle is revealed by the detector. Bars are 280 cm long and with a 4×4 cm^2 cross section oriented along the beam axis all around the inner cryostat surface, installed in the 4.7 cm radial space between the outer surface of COT and the cryostat of the super-conducting solenoid at an average radius of 140 cm, which corresponds to 4.7 ns flight-time for a particle roughly at the speed of light.

Both longitudinal sides of the bars collect the light produced by the charged particles in the scintillator bars into 432 fine-mesh, 19-stage photon-multipliers and measure accurately the timing of the two pulses. The time between the bunch crossing and the scintillation signal in these bars defines the β of the charged particle while the momentum is provided by the tracking system. PID information is available through the combination for a track of TOF and track energy loss measurement from the COT.
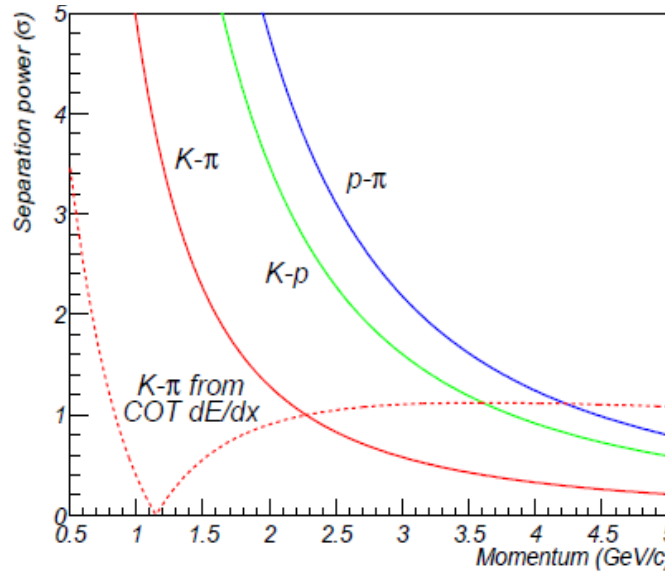
Figure 2.14: Separation power of TOF for different particles at CDF, with dE/dx separation power for kaon and pion from COT superimposed.

The design of the photon-multipliers permits them to maintain an adequate gain even in the 1.4 T magnetic field. The preamplified PMT signals follow two parallel paths: the timing signal is discriminated and digitized, while the charge signal is digitized to be eventually used at trigger level and for subsequent extraction of the offline corrections.

Using the time measurement from this detector and the measured momentum from the COT is possible to infer the particle mass by the formula:

$$m = \frac{p}{c}\sqrt{\frac{c^2 t^2}{L^2} - 1} \qquad (2.5)$$

where p is the momentum measured, L is the path length of the track, and t is the difference between the arrival time of the TOF signal with respect to the the bunch-crossing time.

The expected separation power[11] for the various particle species that is achievable with TOF alone, assuming ≈ 110 ps for the time of flight resolution, as a function of momentum, is shown in Fig. 2.14. For comparison, the expected K/π separation

---

[11]The expected separation power is defined as $\frac{TOF_i(p) - TOF_j(p)}{\sigma_{TOF}}$ where $TOF_i(p) = \frac{L}{c}\sqrt{\frac{m_i c^2}{p^2} + 1}$ is the expected time of fight of the i particle of mass m_i and momentum p. σ_{TOF} is the time of flight resolution.

from the COT dE/dx measurement is also shown in Fig. 2.14 to illustrate the complementary power of COT with respect to the TOF particle identification.

2.5 Calorimeter System

Even if not used in this analysis, the calorimeter system, together with the muon and tracking systems, is one of the main sub-detector apparatus of CDF II detector. A detailed description of this system can be found in [42]. The CDF II calorimetry system has been designed to measure energy and direction of neutral and charged particles leaving the tracking region. In particular, it is devoted to jet reconstruction and it is also used to measure the missing energy associated to neutrinos.
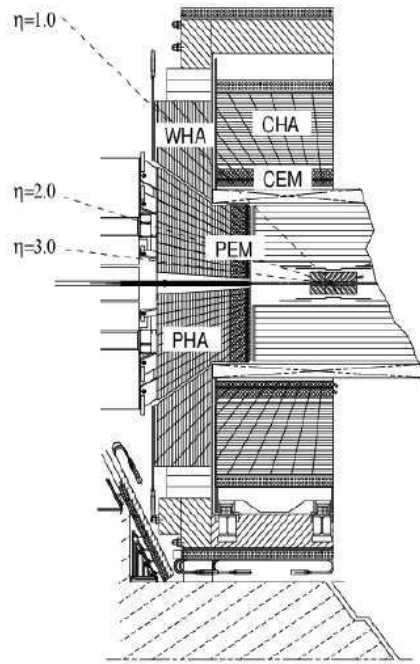

Figure 2.15: Elevation view of the CDF detector showing the components of the CDF calorimeter: CEM, CHA, WHA, PEM and PHA.

Particles hitting the calorimeter can be divided in two classes, according to their main interaction with the matter: electromagnetically interacting particles, such as *electrons* and *photons*, and hadronically interacting particles, such as *mesons* or

baryons produced in hadronization processes. To detect these two classes of particles, two different calorimetric parts have been developed: an inner electromagnetic and an outer hadronic section, providing coverage up to $|\eta| < 3.64$. In order to supply information on particle position, the calorimeter is also segmented in radial sections, called *towers*, projected toward the geometrical center of the detector. Each tower consists of alternating layers of passive material and scintillator tiles. The signal is read out via wavelength shifters (WLS) embedded in the scintillator and the light from WLS is then carried by light guides to photomultiplier tubes. The central sector of the calorimeter, covering the region $|\eta| < 1.1$, was recycled from Run I, while brand new calorimeters (called plug calorimeters) were built up to cover the forward and backward regions. Fig. 2.16(b) shows the plug calorimeter system while Fig. 2.15 shows an elevation view of the components of the CDF calorimeter: the Central Electromagnetic Calorimeter (CEM), the Central Hadronic Calorimeter (CHA), the Wall Hadronic Calorimeter (WHA), the Plug Electronic Calorimeter (PEM), and Plug Hadronic Calorimeter (PHA).

The Central Calorimeter

Apart from upgrades on the readout electronics, needed to follow the increased collision rate, the central calorimeter is almost the same used during Run I. The CEM is segmented in $\Delta\eta \times \Delta\varphi = 0.11 \times 15°$ projective towers consisting of alternate layers of lead and scintillator, while the CHA and CWA, whose geometry tower segmentation matches the CEM one, use iron layers as radiators.

A perspective view of a central electromagnetic calorimeter module (*wedge*) is shown in Fig. 2.16(a), where both the arrangements in projected towers and the light-gatering system are visible. The projective geometry has been used in order to take advantage of the momentum conservation in the transverse plane: before the $p\overline{p}$ collision, the projection in the transverse plane w.r.t. the beam direction of the beam energy is null, therefore this quantity have to be the same also after the collision took place. Thus, for each tower the transverse energy E_T is defined as $E_T = E sin\theta$, where E is the energy detected by the tower and θ is the angle between the beam axis and the tower direction, in the CDF detector coordinates

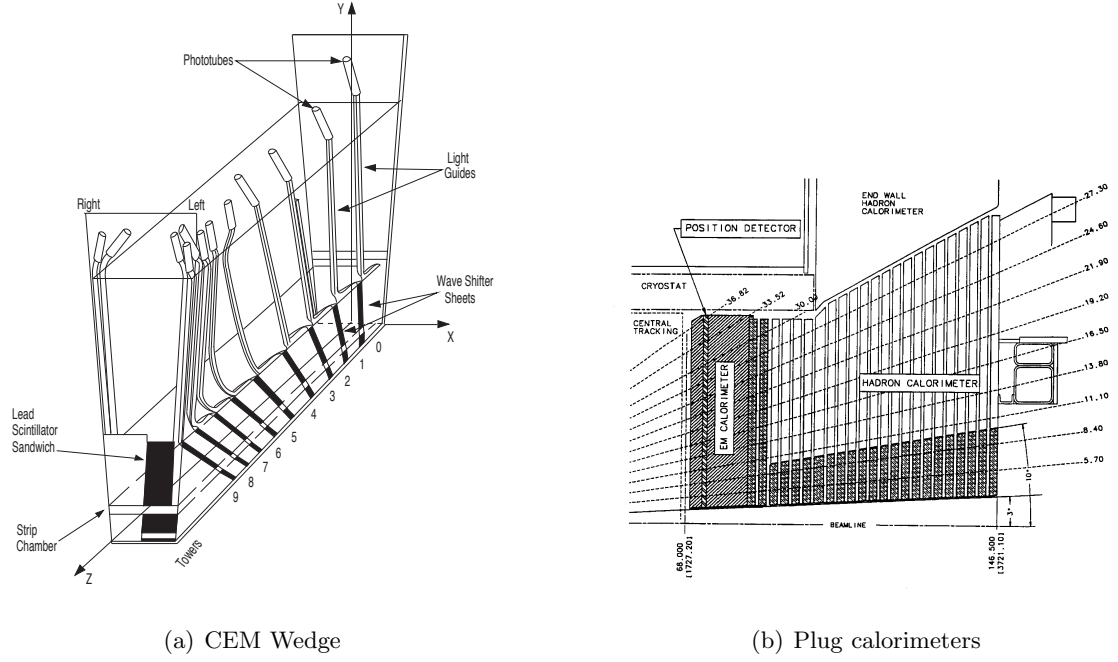(a) CEM Wedge

(b) Plug calorimeters

Figure 2.16: The plot show one azimuthal electromagnetic calorimeter wedge 2.16(a), the second plot shows an elevation view of one quarter of the plug calorimeter 2.16(b).

system. Two position detectors are embedded in each wedge of CEM:

- The **CES**, Central Electromagnetic Strip multi-wire proportional chambers, is a two-dimensional stripwire chamber arranged in correspondence to maximum shower development ($\sim$ 5.9 X_0 [12] It measures the charge deposit of the electromagnetic showers, providing information on their pulse-height and position with a finer azimuthal segmentation than calorimeter towers. This results is an increased purity on electromagnetic object reconstruction.

- The **CPR**, Central Pre-Radiator, consists of two wire camber modules placed immediately in front of the calorimeter. It acts as pre-shower detector by using the tracker and the solenoid coil material as radiators, resulting to be a very useful tool in rejection of electron and photon background.

Calorimeter response is fast enough to match the time requirements imposed by Run II. However, wire chambers associated to CES and CPR may need to be

---

[12]The radiation length X_0 is the scale length describing the high-energy electromagnetic cascades, usually measured in g·cm^{-2}. It is both the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon.

Calorimeter	CEM	CHA	WHA	PEM	PHA
Absorber	Lead	Steel	Steel	Lead	Iron
Segmentation ($\eta \times \varphi$)	0.1×15	0.1×15	0.1×15	$(0.1 \div 0.6) \times (7.5 \div 15)$	$(0.1 \div 0.6) \times (7.5 \div 15)$
Num. Towers ($\eta \times \varphi$)	20×24	9×24	6×24	12×24(48)	11×24(48)
Thickness	18 X_0,1 λ_{int} [13]	4.7 λ_{int}	4.7 λ_{int}	23 X_0,1 λ_{int}	6.8 λ_{int}
Resolution (%)	$14/\sqrt{E_T} \oplus 2$	$50/\sqrt{E_T} \oplus 3$	$75/\sqrt{E_T} \oplus 4$	$16/\sqrt{E} \oplus 1$	$80/\sqrt{E} \oplus 5$

Table 2.4: Summary of the main characteristics of the CDF II calorimeter system.

integrated over several beam crossings.Tab. 2.4 summarizes the basic quantities of calorimeter detectors.

The Plug Calorimeter

The plug calorimeter (see Fig. 2.16(b)), covers the η region from 1.1 to 3.64. The new configuration, based on the same principles as the central calorimeter, allows the detector to operate in the Run II environment and makes experimental data more homogeneous. Both electromagnetic and hadronic sectors are divided in 12 concentric η regions, with $\Delta\eta$ ranging from 0.10 to 0.64, according to increasing pseudo-rapidity, each of them is segmented in 48 or 24 (for $|\eta| < 2.11$ or $|\eta| > 2.11$ respectively) projective towers. The actual size of these towers was chosen so that identification of electron in *b*-jets would be optimized. Projective towers consist in alternating layers of absorbing material (lead and iron for electromagnetic and hadronic sector respectively) and scintillator tiles. The first layer of the electromagnetic tile is thicker (10 mm instead of 6 mm) and made of material with higher photon yield. They act as a pre-shower detector.

2.6 Muon Detectors

The particles produced by the interaction and subsequent decays are absorbed by the system described above with a very high probability. The most common particle that escapes the calorimetric system is the muon. Muons are over 200 times more massive than electrons, so bremsstrahlung radiation, inversely proportional to the mass squared of the incident particle, is suppressed by a factor of $4 \cdot 10^4$ with respect to electrons. They do not interact via strong interaction with nuclei in matter either.

Therefore, a muon with enough energy will pass through the calorimeter systems releasing only a small amount of its energy. The minimum muon energy required to reach the muon detectors, placed radially outside of the calorimeters, is 1.4 GeV.
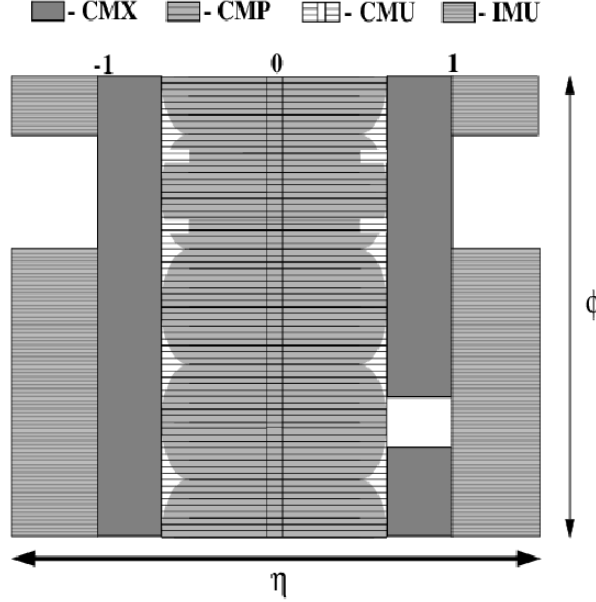

Figure 2.17: Muon detectors coverage in the $\eta - \varphi$ plane.

The muon system is the outermost layer of the CDF II detector and consists on drift cells and scintillation counters which are used to reconstruct segment of track, called *stubs*, from minimum ionizing particles. These stubs are matched using dedicated algorithms with the COT information in order to reconstruct the full trajectory of the muons. Some additional steel shielding layers, in between the chambers and the calorimeters, reduce the probability for other particles to escape the calorimetric system. Four independent systems detect penetrating charged particles in the $|\eta| \lesssim 1.5$ pseudo-rapidity range [56], [57]. The momentum measurement is performed by pointing back the stub to the corresponding track in the COT. Scintillators serve as trigger or trigger veto for muons while the drift chambers measure the φ coordinate using the absolute difference of drift electrons arrival time between two cells, and the z coordinate by charge division.

All type of muon detectors use a single wire, rectangular drift chambers, arranged

in arrays with various azimuthal segmentation and coupled with scintillator counters. The chambers use a 50:50 gas admixture of Argon and Ethane, and operates in proportional regime. The four sub-detector systems are (see Fig. 2.17 for their coverage in the $\eta - \varphi$ plane):

CMU: the Central MUon detector (CMU) detector is located around the central hadronic calorimeter at a radius of 347 cm from the beamline with coverage $0.03 \lesssim |\eta| \lesssim 0.63$. It is segmented into 24 wedges of $15°$. Only $12.6°$ in φ, with a gap of $2.4°$, of each wedge is active, resulting in an overall azimuthal acceptance of 84%. Each wedge is further segmented into three $4.2°$ modules each containing four layers of four drift cells.

CMP: the Central Muon uPgrade (CMP) is the second set of muon drift chambers outside of CMU with an additional 60 cm thick steel absorbers between them. The material further reduces the probability of hadronic punch-through to the CMP. Muons need a transverse momentum of about 2.2 GeV/c to reach the CMP. The CMP system is arranged in a box shape of similar acceptance as the CMU and conventionally serves as a confirmation of CMU for higher momentum muons. The CMP and CMU have a large overlap in coverage and are often used together. CMP helps to cover CMU φ gaps and the CMU covers the CMP η gaps. Muon candidates which have both CMU and CMP stubs are the less contaminated by fake muons.

CMX: the Central Muon eXtension (CMX) consists of drift tubes and scintillation counters (CSX) assembled in conically arranged sections. The CMX extends the pseudo-rapidity coverage to $0.6 \lesssim |\eta| \lesssim 1$. There are 8 layers of drift chambers in total with a small stereo angle between layers.

IMU: the Intermediate MUon system (IMU) extends the pseudo-rapidity coverage even further to $1.0 \lesssim |\eta| \lesssim 1.5$. The IMU is mounted on the toroid magnets which provide shielding and consists of the Barrel MUon chambers (BMU), Barrel Scintillation counters (BSU) and Toroid Scintillation counters (TSU).

2.7   Cherenkov Luminosity Counters (CLC) and Measurement of the Luminosity

The main purpose the CLC (see Fig. 2.18) is to measure the instantaneous luminosity $\mathcal{L}$ in the BØ interaction point. The relation used is $\bar{N} \times f_{b.c.} = \sigma_{p\bar{p}} \times \varepsilon \times \mathcal{L}$, where $\bar{N}$ is the number of interaction for bunch crossing (BC), $f_{b.c.}$ the BC frequency, on average 2.5 MHz for 36×36 bunch operations, $\sigma_{p\bar{p}}$ is the inelastic cross section[14], and ε is the detector efficiency. The inelastic cross section is obtained by extrapolating the combined results for the inelastic $p\overline{p}$ cross section of CDF at $\sqrt{s} = 1.8$ TeV and measurements at $\sqrt{s} = 1.96$ TeV [58] at the Run II energy. The global uncertainty on the luminosity is $\approx 5.6\%$. This detector covers the $3.7 \lesssim |\eta| \lesssim 4.7$ range,
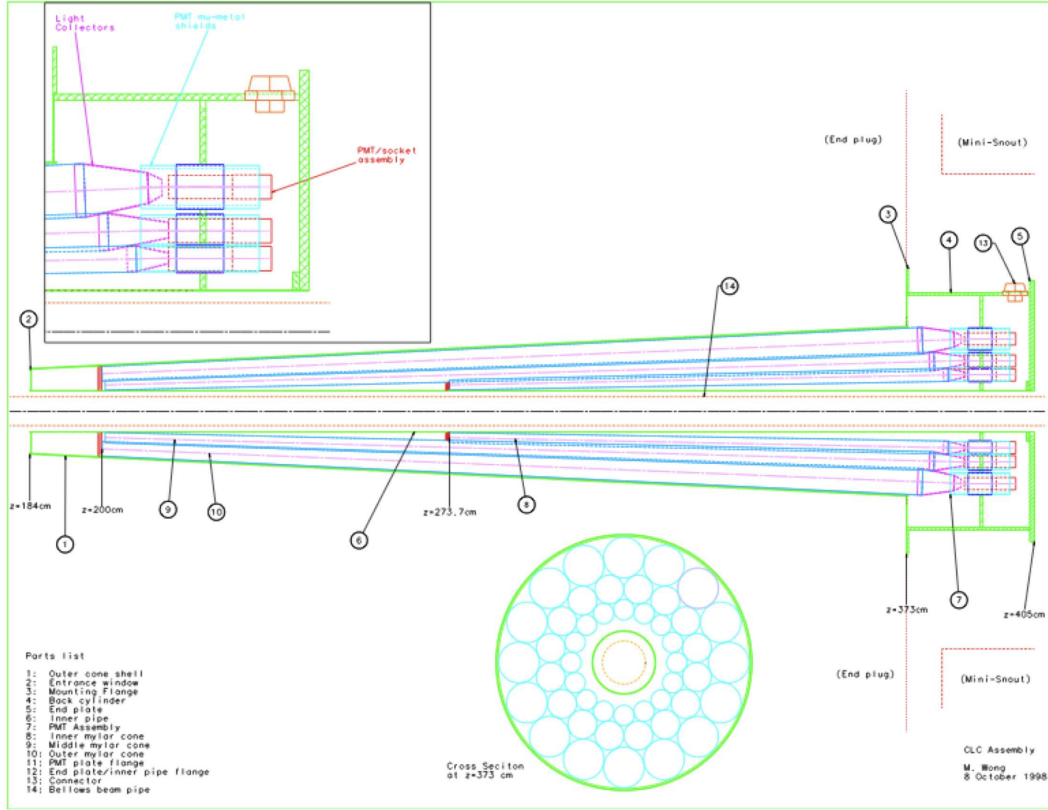

Figure 2.18: Longitudinal section of the CLC system forward.

with two symmetrical detector placed in the forward and in the backward regions. It is composed by long Cherenkov detectors, $100 - 108$ cm, with a conical shape,

---

[14]The proton-antiproton inelastic cross section at the Tevatron is $\sigma_{p\bar{p}} = 59.3$ mb.

filled with Isobutane that guarantees high refraction index and good transparency for ultraviolet photons. The light emitted are collected in a PMT shielded by the solenoidal magnetic field. The luminosity measured by the CLC is used to monitor the Tevatron performance.

2.8 Trigger and DAQ

At the typical Tevatron instantaneous luminosity $\mathcal{L} \approx 4 \cdot 10^{32}$ cm^{-2}s^{-1}, and with an inelastic $p\bar{p}$ cross section of $\sigma_{p\bar{p}} \approx 60$ mb, approximately $2.5 \cdot 10^7$ inelastic collisions per second occur, corresponding to one inelastic $p\overline{p}$ interaction per bunch-crossing on average. Since the read-out of the entire detector needs about 2 ms on average, after the acquisition of one event, another approximately $5,000$ interactions would remain unrecorded. When an event recording is prevented because the system is busy with a different event or a different task, this is called *dead-time*.

The average size of information associated to each event is 140 Kb. Even in case of deadtime-less read-out of the detector, in order to record all events, an approximate throughput and storage rate of 350 Gb/s would be needed, largely beyond the possibility of currently available technology[15]

The read-out system has to reduce the 2.3 MHz interaction-rate to the 100 Hz storage rate attainable at CDF. The challenge for the whole system is to cut-off events that don't have the minimal requirements to be reconstructed, focusing the acquisition system on the interesting processes. The Data Acquisition Quality (DAQ) is segmented in three levels (see Fig. 2.19). Each level receives the accepted event from the previous one, and verifies if it passes particular requirements.

Prior to any trigger level, the bunched structure of the beam is exploited to reject cosmic-ray events by gating the front-end electronics of all sub-detectors in correspondence of the bunch crossing. The front-end electronics of each sub-detector, packaged in VME modules hosted in about 120 crates, has a 42 cells deep pipeline synchronized with the Tevatron clock-cycle set to 132 ns. The Tevatron clock picks up a timing marker from the synchrotron RF and forwards this bunch-crossing signal

---

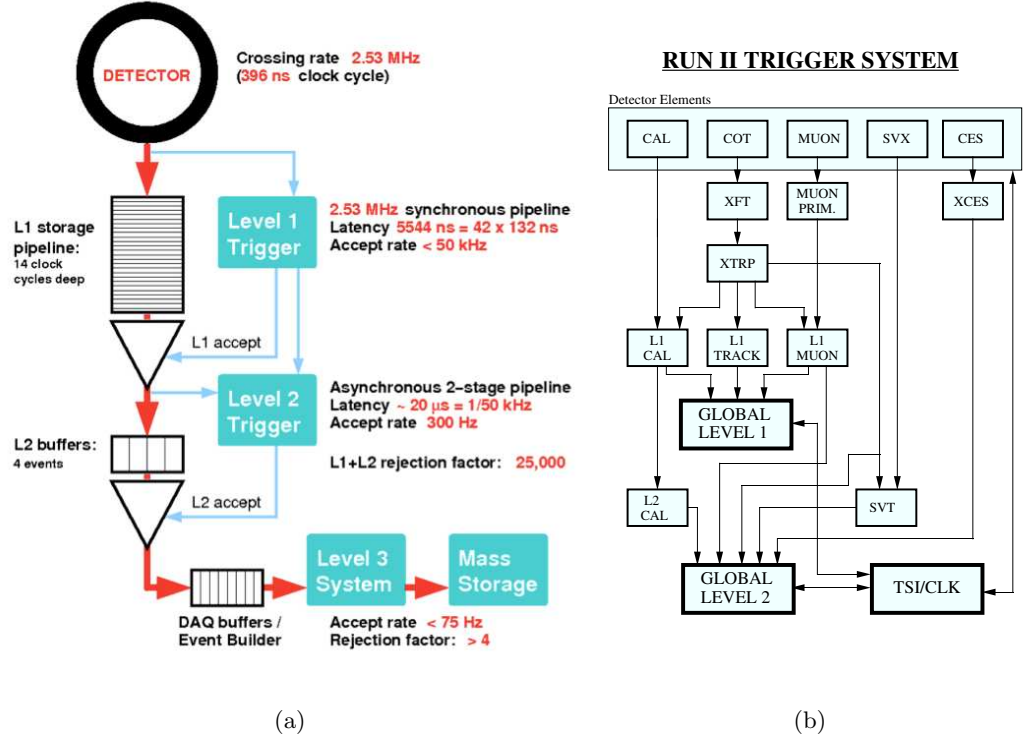[15]The maximum current storage rate is approximately 250 Kb/s.

(a)

(b)

Figure 2.19: Block diagram showing the global trigger **2.19(a)** and the data flow for the L1 and L2 systems at CDF II **2.19(b)**.

to the trigger and to the front-end electronics. Since the inter-bunch time is 396 ns, three times the Tevatron clock-cycle, the pipeline can collect data corresponding to a maximum of 14 bunch crossings. The pipeline depth gives the amount of time that Level-1 (L1) trigger has to decide to accept or reject an event otherwise the buffer content is overwritten: 396 ns · 14 = 5.5 μs. An event accepted by the L1 is then passed to the Level-2 (L2) buffer, where the number of buffers in the pipeline is 4, that gives 5.5 μs ·4 = 22 μs. This means that if an event is accepted by the L1 and the L2 doesn't have a free buffer, deadtime will incur. Level-3 (L3) is composed by a computer farm, the L2 output rate is low enough to avoid in general deadtime problem in the connection between L2 and L3. The following description will emphasize the aspect of the CDF Trigger that are related with the selection of rare events including *b*-hadrons with high purity.

2.8.1 Level-1

L1 is a synchronous system of custom designed hardware which reconstructs information from three parallel streams: the calorimeter (total energy and single tower information), the COT (only 4 axial superlayers are used for two-dimensional tracks), and the muon systems (stubs in the CMU, CMP and CMX).

Two-dimensional tracks in the $r - \varphi$ plane of the COT are identified by the eXtremely Fast Tracker XFT (see Sect. 2.9.1). It sends the track information to the extrapolation unit (XTRP) and then trigger objects are formed using combinations of primitives and cuts on those primitives. Calorimeter based objects (electrons, jets, photons) are formed in the L1CAL, muons are formed in the L1MUON and hadronic B are formed in the L1TRACK.

A decision stage combines the information from these low-resolution physic objects, called *primitives*, into more sophisticated objects.

The accepted events are buffered for L2 analysis.

2.8.2 Level-2

The L2 is an asynchronous system of custom-designed hardware which processes events accepted by the L1.

Additional information from the shower-maximum strip chambers in the central calorimeter and the axial hits in the SVX II are combined with the L1 primitives to produce L2 primitives. A simplified energy-clustering is done in the calorimeters, merging the energies in adjacent towers to the energy of a seed tower above threshold. L1 track primitives matched with consistent shower-maximum clusters provided refined electron candidates whose azimuthal position is known within $2°$ accuracy.

Information from the $r - \varphi$ sides of the SVX II is combined with L1 tracks primitives to form two-dimensional tracks with resolution similar to the offline one by the silicon-based trigger system, (see Sec. 2.9.2). SVT uses SVX $r - \varphi$ hits to extend XFT track primitives inside the SVX volume, closer to beamline. The SVT improves the XFT φ_0 and p_T resolutions and adds the measurement of the impact parameter d_0 (original XFT track primitives are beamline constrained). Acting into

the impact parameter, SVT is a very useful handle in order to select decay modes of heavy *b*-hadrons into charged prongs. As shown in Fig. 2.20, the impact parameter
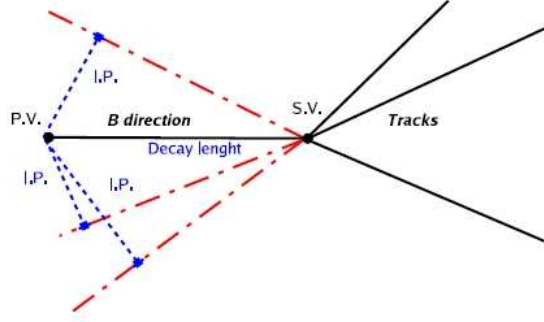

Figure 2.20: Schematic chart showing the correlation between the tracks impact parameter (I.P.) and the decay length, in the transverse plane, of a hypothetical *b*-hadron decay.

(IP) of the decay products is strongly related to the decay length of the mother *b*-hadron, therefore a selection based on the tracks impact parameter turns directly into a proper time requirement. This innovative system is the core of all the trigger systems for *B* physics, and will be described in further details in Sec. 2.9.

Finally, an array of programmable processors makes the trigger decision, while the L2 objects relative to the following event accepted at L1 are already being reconstructed. The L2 output rate is around 900 Hz.

Fig. 2.19(b) shows the block diagram of the CDF II trigger system with the L1 and L2 subsystems along with their interconnections.

2.8.3 Level-3

After an event is accepted at L2, it has to be read out completely. The digitized output relative to the L2-accepted event arrives fragmented from all sub-detectors via optical fibers. It is collected by the Event Builder, a custom hardware switch that

arranges it in the proper order and assembles in the *event record*. It is then transfered to one of the nodes in the L3 farm, consisting of 300 commercial CPUs, organized in a modular and paralleled structure of 16 sub-systems [59], each composed by 12 to 16 processor nodes.

The L3 trigger reconstructs the event following given algorithms, taking advantage of the full detector information and improved resolution not available to the lower rigger levels. This includes a full three-dimensional track reconstruction and tight matching of tracks to calorimeter and muon-system information. Events that satisfy the L3 trigger requirements are then transfered onward to the Consumer Server/Data Logger (CSL) system for storage first on disk and later on tape. The average processing time per event in L3 is of the order of a few seconds. The L3 leads to a further reduction in the output rate, roughly 75 Hz.

A set of requirements that an event has to fulfill at L1, L2 and L3 constitutes a trigger path. The CDF II trigger system implements about 200 trigger paths. An event is accepted if it passes the requirements of any one of these paths and, depending of the particular trigger path, it will be stored in a trigger dataset. A complete description of the different datasets of the CDF Run II can be found in [60].

Another important feature of the trigger system of CDF is that L1 and L2 accepted events can be pre-scaled. This means that only a fraction of the events that fulfill the trigger requirements are actually accepted. Even if this implies loosing potentially useful events, it becomes necessary at high luminosity. Given the continuous improving performance of the Tevatron, pre-scaling trigger has become common practice in the last years. Moreover, the trigger system allows for dynamic pre-scaling of trigger acceptance, meaning that the scaling factor varies with the instantaneous luminosity, so the output bandwidth is maximally utilized. If an event satisfies the L3 requirements, the corresponding event record is transferred to mass storage at a maximum rate of 20 Mbyte/s.

The L3 decision is made after the full reconstruction of the event is completed and the integrity of its data is checked, a process that takes a few milliseconds. A fraction of the output is monitored in real time to search for detector malfunctions, to calibrate and to graphically display events.

2.9 The Track Trigger

The goal of the trigger is to collect the largest number of rare *b*-hadron decays in the limited amount of available bandwidth.

Also it is important to have a good background rejection, with a reasonable efficiency on the interesting signal. This can be achieved reconstructing the track parameters in real-time. In the next sub-sections will be described the two processor that are doing this task at level 1, XFT (see Sec. 2.9.1), and at level 2, SVT (Sec. 2.9.2).

2.9.1 The COT track-processor: XFT

The COT is connected to the trigger processor XFT that uses pattern matching in the $r - \varphi$ plane to first identify short segments of tracks and to link them into full-length tracks, in time with the L1 decision.

The XFT tracks are also extrapolated to the electromagnetic calorimeter and muon chambers to generate trigger electron and muon candidates. The role of the XFT in the trigger is to reduce the raw collision rate (1.7 MHz) to a maximum of about 30 KHz of interesting physics events which can be processed by the L2 trigger.

The XFT uses hit data from the 4 axial superlayers of the chamber, arranged in cells of 12 wires each. A charged particle passing through an axial layer generates a characteristic pattern of 12 hits, one per wire, with a characteristic timing. Track identification is performed in two steps by the Finder boards, which search for track segments in each of the 4 axial layers of the chamber, and by the Linker boards, which search for 4/4 matches among segments in the 4 layers, consistent with a track exiting from the interaction point. If a coincidence between segments crossing four super-layers is found, two-dimensional XFT-tracks are reconstructed by linking the segments. The segments are compared with a set of about 2,400 predetermined patterns corresponding to all possible tracks with $p_T \geq 1.5$ GeV/c originating from the beam line. The comparison proceeds in parallel each of the 288 azimuthal 1.25° sectors in which XFT logically divides the chamber. If no track is found using all four super-layers, then the best track found in the innermost three super-layers is

taken as output. The track-finding efficiency and the fake rate with respect to the offline tracks depends of the instantaneous luminosity, and were measured to be $\varepsilon \approx 96\%$, and 3% respectively, for tracks with $p_T \geq 1.5$ GeV/c at $\mathcal{L} \simeq 10^{31}$ cm^{-2}s^{-1}. The observed momentum resolution is $\sigma_{p_T}/p_T^2 = 0.017$ [GeV/c]$^{-1}$, and the azimuthal resolution is $\sigma_{\varphi_6} = 0.3°$, where φ_6 is the azimuthal position at the sixth COT super-layer, located at 106 cm radius from the beam line. The reconstructed tracks are reported to the XTRP and a copy of them is preserved to be used in the L2.

2.9.2 The Online Silicon Vertex Tracker (SVT)

The Online Silicon Vertex Tracker (SVT) (see Fig. 2.21) is part of the L2 trigger.
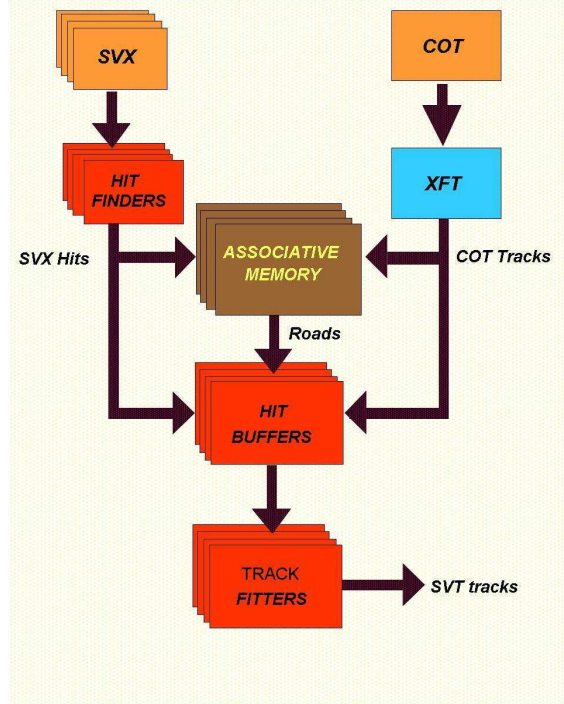

Figure 2.21: SVT architecture.

It receives the list of the COT tracks reconstructed by the XFT processor (for each track the curvature C and the azimuthal angle φ are measured) and the digitized pulse heights on the silicon layers (10^5 channels). The SVT links the XFT tracks to the silicon hits and reconstructs tracks in the transverse plane with offline-like quality. The resolution of the SVT is $\delta\varphi \simeq 1.0$ mrad, $\delta p_T \simeq 0.003 \cdot p_T^2$ GeV/c

and $\delta d_0 \simeq 35 \; \mu m$, where d_0 is the track impact parameter, the radial distance of closest approach of the particle trajectory helix to the z-axis of the CDF reference system. The SVT efficiency, defined as the ratio between the number of tracks reconstructed by SVT and all XFT-matched offline silicon tracks that are of physics analysis quality, is higher than 85%. By providing a precise measurement of the impact parameter of charged particle tracks, SVT allows on-line triggering on events containing long lived particles. Hadrons with b quark in particular have a decay length of the order of 500 μm and tracks which come out of the b-hadrons decay vertices have an impact parameter on average grater than 100 μm.

The SVT has a widely parallelized design: it is made of 12 identical slices (*wedges*) working in parallel. Each wedge receives and processes data only from tracks in the plane perpendicular to the beamline (stereo info from SVXII is dropped) and only with p_T above 2 GeV/c. The tracking process is performed in two steps:

- **Pattern recognition**: candidate tracks are searched among a list of precalculated low resolution patterns (*roads*);

- **Track fitting**: a full resolution fit of the hit coordinates found within each road is performed using a linearized algorithm.

The pattern recognition step is performed in a completely parallel way by the AM system which uses full custom VLSI chips (AMchips [61]). The AM system compares all the silicon clusters and XFT tracks with the set of precalculated patterns. Each SVT wedge uses 32,000 patterns which cover more than 95% of the phase space for $p_T \geq 2$ GeV/c.

The track fitting method is based on linear approximations and principal components analysis [62].

Fig. 2.22 shows the SVT track impact parameter resolution for tracks with $p_T > 2$ GeV/c. The width of the Gaussian fit for the distribution in Fig. 2.22 is 55 μm. This is a combination of the intrinsic SVT impact parameter resolution, and the transverse size of the beam line: $\sigma_{fit} = \sigma_{SVT} \oplus \sigma_{beam}$, where σ_{beam} is about 30 μm. Therefore, the intrinsic SVT resolution is about 35 μm. SVT allows, for the first time at a hadron collider, to trigger directly on hadronic b decays with charged
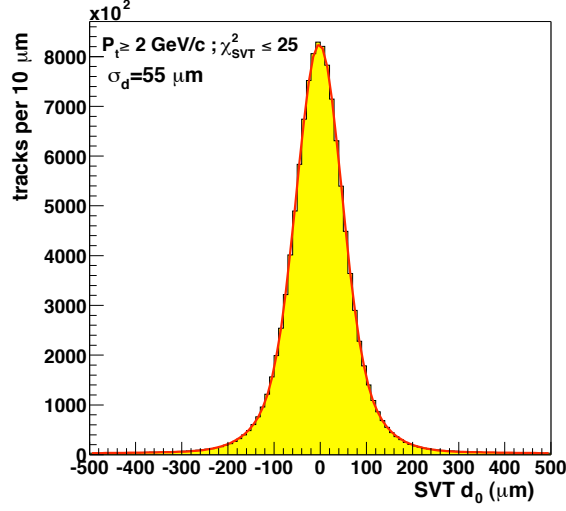
prongs.

Figure 2.22: The plot shows the distribution of the impact parameter as measured by the SVT processor.

2.9.3 The Two Track Trigger Path (TTT)

In the following we will describe the TTT trigger path used for the analysis described in this Thesis. The term Two Track Trigger (TTT) is used within CDF to indicate two kind of triggers that require at least two charged tracks in the event, with some kinematical constraint that we are going to describe in the following. These triggers are able to extract fully hadronic decays from a large background of tracks, just using the tracks reconstructed by SVT. The two triggers differ by the fact that one is meant to collect two-body decays, like $B_s^0 \to h^+ h'^-$ (B_PIPI), while the other is sensitive to multi-body decays, like the one we are looking for, $B^- \to D\pi^-$ (B_CHARM).

We are going to describe in more details the latter, since it has been used in this Thesis. Table 2.5 reports the three main variants of this trigger, according to the requirement in the minimum transverse momentum: low$-p_T$, medium$-p_T$ and high$-p_T$.

The impact parameter (d_V) and transverse decay length (L_{xy}) of the B candidate vertex are calculated starting from the parameters of the two SVT tracks using the

Version	Level 1	Level 2	Level3
High$-p_T$	$p_T > 2.5$ GeV/c	$p_T > 2.5$ GeV/c	$p_T > 2.5$ GeV/c
	Opposite charge	Opposite charge	Opposite charge
	$\Delta\phi_6 < 90°$	$2° < \Delta\phi_0 < 90°$	$2° < \Delta\phi_0 < 90°$
	$\sum_{p_T} > 6.5$ GeV/c	$\sum_{p_T} > 6.5$ GeV/c	$\sum_{p_T} > 6.5$ GeV/c
		$0.1 < d_0 < 1$ mm	$0.1 < d_0 < 1$ mm
		$L_{xy} < 200$ μm	$L_{xy} < 200$ μm
			$|\Delta z_0| < 5$ cm
			$|\eta| < 1.2$
Medium$-p_T$	$p_T > 2$ GeV/c	$p_T > 2$ GeV/c	$p_T > 2$ GeV/c
	Opposite charge	Opposite charge	Opposite charge
	$\Delta\phi_6 < 90°$	$2° < \Delta\phi_0 < 90°$	$2° < \Delta\phi_0 < 90°$
	$\sum_{p_T} > 5.5$ GeV/c	$\sum_{p_T} > 5.5$ GeV/c	$\sum_{p_T} > 5.5$ GeV/c
		$0.1 < d_0 < 1$ mm	$0.1 < d_0 < 1$ mm
		$L_{xy} < 200$ μm	$L_{xy} < 200$ μm
			$|\Delta z_0| < 5$ cm
			$|\eta| < 1.2$
High$-p_T$	$p_T > 2$ GeV/c	$p_T > 2$ GeV/c	$p_T > 2$ GeV/c
	Opposite charge	Opposite charge	Opposite charge
	$\Delta\phi_6 < 90°$	$2° < \Delta\phi_0 < 90°$	$2° < \Delta\phi_0 < 90°$
	$\sum_{p_T} > 4$ GeV/c	$\sum_{p_T} > 4$ GeV/c	$\sum_{p_T} > 4$ GeV/c
		$0.1 < d_0 < 1$ mm	$0.1 < d_0 < 1$ mm
		$L_{xy} < 200$ μm	$L_{xy} < 200$ μm
			$|\Delta z_0| < 5$ cm
			$|\eta| < 1.2$

Table 2.5: Selections of the three versions of the B_CHARM trigger. Criteria refer to track-pairs. The p_T, d_0 and η requirements are applied to both tracks. The $\sum_{p_T}$ refers to the scalar sum of the p_T of the two tracks. ϕ_6 is the azimuthal angle of the tracks measured at the COT superlayer 6. It is related to ϕ_0 by the relation $\phi_6 = \phi_0 + arcsin(R \cdot C)$, where $R = 105.575$ cm and C is the track's curvature.

following equations:

$$p_x = p_{T1} \cos \phi_1 + p_{T2} \cos \phi_2 \quad p_y = p_{T1} \sin \phi_1 + p_{T2} \sin \phi_2.$$

The flight direction of the B candidate vertex is then given by:

$$\cos \phi_V = \frac{p_x}{p_T} \quad \sin \phi_V = \frac{p_y}{p_T},$$

where

$$p_T = \sqrt{p_x^2 + p_y^2} \; .$$

The coordinates of the B candidate vertex are:

$$x_V = \frac{d_1 \cos \phi_2 - d_2 \cos \phi_1}{\cos \phi_1 \sin \phi_2 - \cos \phi_2 \sin \phi_1} \qquad y_V = \frac{d_1 \sin \phi_2 - d_2 \sin \phi_1}{\cos \phi_1 \sin \phi_2 - \cos \phi_2 \sin \phi_1} \qquad (2.6)$$

from which:

$$L_{xy} = \frac{x_V \cdot p_x + y_V \cdot p_y}{p_T} \qquad d_V = x_V \sin \phi_V - y_V \cos \phi_V \qquad (2.7)$$

2.10 CDF MC simulation of detector and trigger

In a modern high energy physics experiment is important to have a precise simulation of the detector response and ability to reconstruct energy and momentum of the particles. The geometry of a detector is quite complex, so in order to reduce the uncertainties over the detector capabilities, the use of a detailed Monte Carlo simulation (MC) which simulates the efficiency of reconstruct a decay channel, the geometrical acceptances, and other similar tasks, is mandatory.

To generate a B-hadron decay sample there are three different steps: the first step is the B-hadron generation, then the hadron is forced to decay into a particular final state and finally the decay products are propagated within the CDF II detector. The hadron generation and the decay of the sample used in this Thesis are generated using the "BGenerator" algorithm, which needs, as input information, the joint distribution of transverse momentum and pseudo-rapidity for the different b-hadrons.

Once the physics of the interaction and the generated particles are available, they are propagated within a simulation of the detector in order to reproduce their interactions with the materials and the different signals they produce in all the sub-

detectors. In the standard CDF simulation, the detector geometry and material are modeled using the version 3 of the GEANT package [63], tuned using data from test-beams and actual collisions. GEANT receives as input the positions, the four-momenta, and the identities of all particles produced by the simulated collisions that have long enough lifetimes to exit the beam pipe. It simulates their passage in the detector, modeling their interactions (*bremmstrhalung*, multiple scattering, nuclear interactions, photon conversions, etc.) and the consequent generation of signals on a channel by channel basis.

The actual trigger logic is simulated for all digital parts of the trigger. The output of the simulated data has the same format of the collision data, allowing their analysis with the same reconstruction programs used.

The detector and trigger configuration underwent several changes during data-taking. For a more detailed simulation of the actual experimental conditions, the simulation has been interfaced with the offline database that reports, on a run-by-run basis, all known changes in configuration (position and slope of the beam line, relative mis-alignments between sub-detectors, trigger-table used, set of SVT parameters) and local or temporary inefficiencies of the silicon tracker (active coverage, noisy channels, etc.). This allows simulating the detailed configuration of any set of real runs, to match the distribution of real data in any given sample to very high precision.

<div align="right">

3

</div>

Selection of the suppressed signals
$B^- \rightarrow D(\rightarrow K^+\pi^-)h^-$

In this chapter we will describe the procedure to select the suppressed signals $B^- \rightarrow D(\rightarrow K^+\pi^-)h^-$. After the on-line trigger selection, an off-line selection has been performed to see a clear peak of favored $B^- \rightarrow D\pi^-$ events. Then we chose a set of requirements that maximizes the probability of detecting the suppressed $B^- \rightarrow D\pi^-$ peak.

3.1 Data sample

The technique for the measurement of the CKM angle γ proposed by Atwood, Danietz and Soni (see Chap. 1.5.1) is centered on the interference between these two decay channels[1]:

- color-favored $B^- \rightarrow D^0 K^-$ followed by the doubly Cabibbo suppressed $D^0 \rightarrow K^+\pi^-$,

- color-suppressed $B^- \rightarrow \overline{D}^0 K^-$ followed by the Cabibbo favored $\overline{D}^0 \rightarrow K^+\pi^-$.

The internal D^0 and $\overline{D}^0$ states are indistinguishable and we can only reconstruct the final state $[K^+\pi^-]_D K^-$ and measure the direct CP- violating asymmetry. The

---

[1]We imply the charge conjugate state throughout this thesis, except in formulas and sentences where both are mentioned explicitly.

two virtual decay chains make up a single sample that in the following we will call "suppressed" (sup) for brevity.

A similar decay, but less suppressed, is the $B^- \to D(\to K^-\pi^+)K^-$, which is produced by two decay chains:

- color-favored $B^- \to D^0 K^-$ followed by the Cabibbo favored $D^0 \to K^-\pi^+$,

- color-suppressed $B^- \to \overline{D}^0 K^-$ followed by the doubly Cabibbo suppressed $\overline{D}^0 \to K^-\pi^+$.

However the second possibility is doubly suppressed and its contribution is negligible compared with the first.

The $B^- \to D(\to K^-\pi^+)K^-$ mode is clearly a source of background for the suppressed mode $B^- \to D(\to K^+\pi^-)K^-$, the only difference being the identity of the decay products of D: $K^-\pi^+$ in the first case and $K^+\pi^-$ for the second[2]. However its abundance can also be exploited as an advantage using the "favored" (fav) $B^- \to D(\to K^-\pi^+)K^-$ mode to normalize the branching ratio of the suppressed signal.

Another background of the same kind comes from the favored $B^- \to D(\to K^-\pi^+)\pi^-$. Again it can be used to normalize the suppressed $B^- \to D(\to K^+\pi^-)\pi^-$ and to measure the ADS observables also for the pion mode (as explained in Sect. 1.5.1).

For these reasons we need to reconstruct the suppressed $B^- \to Dh^-$ modes (where $h = \pi$ or K), as well as the favored $B^- \to Dh^-$ modes through a general selection of B candidates composed by a D plus a track.

Moreover the favored sample has the same topological features of the suppressed and this allows us to perform a full data-driven analysis, taking the favored sample as a model for the suppressed. In particular we will use the $B^- \to D(\to K^-\pi^+)\pi^-$ signal for this purpose.

The most important kinematic observables describing these decays are (Fig. 3.1):

- The **primary vertex** PV is the space-point of the $p\bar{p}$ collision, where a b-quark pair is produced and quickly hadronizes to a B-hadron pair.

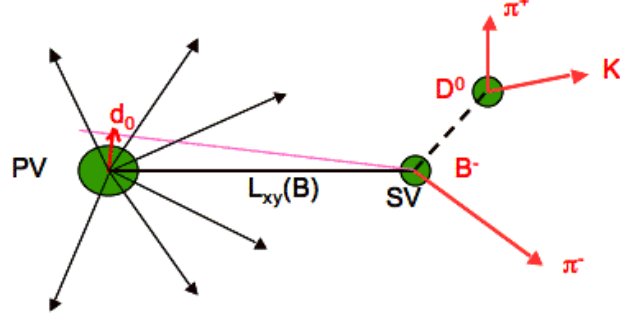[2]With the symbol D we mean both D^0 and $\overline{D}^0$, except in formulas and sentences where they are mentioned explicitly.

Figure 3.1: Schematic diagram of the $B^- \to D(\to K^- \pi^+)\pi^-$ decay channel. Some of the most important kinematic observables are reported.

- The **secondary vertex** SV is the space-point where the long-lived particle decays. Its displacement with respect to the primary vertex is defined as $\vec{x}_v = \vec{\beta}\gamma ct$, where $\vec{\beta}\gamma = \vec{p}/m$ is the ratio between the momentum and the mass of the particle, which decays at a time t after its production. The mean value of t is the mean time life τ of the particle. The large $c\tau$ of the B meson, about 491 μm, can be used to discriminate events from B mesons to background.

- The **transverse decay length** L_{xy} is the signed distance between the primary and the secondary vertices projected onto the transverse momentum vector $\vec{p}_T$ of the decaying particle

$$L_{xy} \equiv \frac{\vec{p_T} \cdot \vec{x}_v}{p_T}$$

  The transverse displacement of the secondary vertex $\vec{x}_v$ may be not collinear with $\vec{p}_T$, the transverse decay length is usually preferred to it. This quantity is typically positive for true long-lived decays, while fake secondary vertex can have negative values or equal to its resolution.

- The **impact parameter** d_0 is signed distance of the closest approach between the flight path of a particle and the beam line in the transverse plane

$$d_0 \equiv \frac{\hat{z} \cdot (\vec{p}_T \wedge \vec{x}_v)}{p_T}$$

  where the scalar product with the unit vector pointing toward the proton direction $\hat{z}$ determines its sign and the symbol $\wedge$ indicates the vector product.

The impact parameter is typically different from zero for products of long-lived decays.

3.2   Reconstruction of the $B^- \to Dh^-$ sample

Data used for this analysis has been collected with the upgraded CDF II detector, between April 2001 and July 2010 by the trigger path B_CHARM. We used events collected in runs where the following systems were declared good by the CDF Data Quality Monitoring Group: SVX, COT, CLC and all the trigger levels. The final sample corresponds to an integrated luminosity of 7 fb^{-1}.

We begin the reconstruction of B candidates by requiring at least two SVT tracks, matching the *Medium-p_T* selection of the B_CHARM selection criteria, listed in Section 2.9.3.

The events selected by the trigger are further discriminated, searching for a pair of oppositely charged tracks, compatible with a two-body D decay. We require that the invariant mass of the two tracks, obtained with the tracks momentum and with a-priori assignment of the kaon and pion masses to them, is compatible with the nominal D mass (~ 1.865 GeV/c^2), in the range $[1.770, 1.970]$ GeV/c^2.

The charmed candidate is then combined with an additional charged track in the event with $p_T > 0.4$ GeV/c, to form the B candidate in the secondary vertex. Events with invariant mass in the interval $[4.4, 6.6]$ GeV/c^2 are selected.

Both the D and the B candidate vertexes are obtained constraining the tracks to belong to a common vertex. We used the CTVMFT algorithm [64], which determines the decay vertex of the track minimizing the distance between the constrained vertex and the original unconstrained one. The method takes into account also the correlations of the tracks parameters evaluated at the distance of closest approach to the z axis.

The parameters of the tracks are refit using the additional information on its position and the reconstructed momentum of the B candidate is constrained to point back to the primary $p\overline{p}$ interaction vertex. A constraint on the mass on the charmed candidate mass is also applied, determining the momenta of the daughter tracks.

We divided the events into two non-overlapping samples, according to the relative charge of the B candidate with the decay product of the D that has been categorized as the kaon. We classify events where the track from B has the same charge sign of the kaon from D in the "favored" sample, and events where the track from B has the same charge sign of the pion from D in the "suppressed" sample.

To ensure no overlap between the two samples, we veto the alternate mass configuration, where the identity of the decay products of the D is swapped. In each sample and for each event we evaluate the "correct"[3] (called "$K\pi$") and the "alternate" (πK), where the D decay products mass assumption is swapped with respect to the "correct" one. In Fig. 3.2 we can see the distribution of the invariant mass of D candidate in the $K\pi$ hypothesis vs πK hypothesis. We veto the intersecting region where the πK mass is compatible with $[1.8245, 1.9045]$ GeV/c^2.
The remaining small amount of events with an incorrect identification of D decay products is accounted as part of the inclusive background $B^- \to D(\to X)\pi^-$, where X are modes other than $K\pi$ (see Sect. 4.2).

Since one of the observable to be measured is the ratio between suppressed and favored events, an important feature of this analysis is to perform symmetric selections on both samples, especially to avoid to correct for selection efficiency in the final observables. For this reason the veto on the "swapped mass" of the D is symmetrically performed on both samples, as well as the following cuts we are going to describe. Indeed, the strict similarity between favored and suppressed signals is a strong point of this analysis, that we are going to extensively exploit.

3.2.1 Basic requirements

The following baseline requirements are applied to both suppressed and favored samples:

- transverse B decay length significance: $\dfrac{L_{xy}(B)}{\sigma_{L_{xy}(B)}} \geq 8$;

- B decay length error: $\sigma_{L_{xy}(B)} \leq 0.01$;

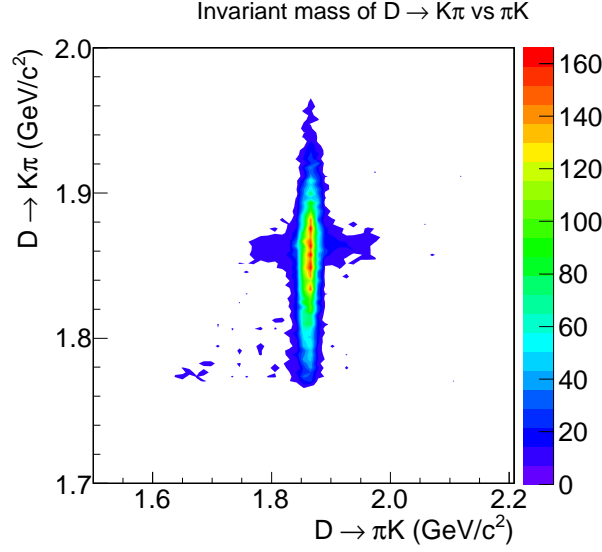---

[3]The term "correct" is used as a convention.

Figure 3.2: Invariant mass distribution of D candidate in the two $K\pi$ vs πK ("swapped") mass assumption of the decay products of the D (suppressed sample).

- B impact parameter: $|d_0(B)| \leq 0.008$ cm;

- B transverse momentum: $p_T(B) \geq 5$ GeV/c^2;

- D decay length measured with respect to the B decay vertex: $L_{xy}(D)_B \geq -0.015$ cm;

- $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the track from B and the D: $\Delta R \leq 2$;

- Eta of the track from B: $|\eta| \leq 1.$;

- Eta of the tracks from D: $|\eta| \leq 1.$;

- transverse momentum of the track from B: $p_T \geq 1$ GeV/c^2;

where the transverse B decay length significance is the ratio between the transverse decay length and its error, ΔR is the distance in the $\eta - \phi$ space between the D and the track from B.

Events where tracks from real B decays are incorrectly labeled as D decay products are a source of background. We veto the invariant mass formed by the track from the B candidate and the oppositely charged track from the D candidate, requiring it to be incompatible with the D meson mass, using the same range as

the first veto ($[1.8245, 1.9045]$ GeV/c^2). The cut is symmetrically applied on both samples.

A further suppression of this background is achieved by requiring that the transverse distance between B and D decay vertex $L_{xy}(D)_B$ is greater than 100 μm. This has the additional effect of reducing contamination from non-resonant three-body decays of the type $B^+ \to h^+ h^- h^+$, in which all tracks come from a common decay vertex, and where h indicates either K or π.

As can be seen in Fig. 3.3, three-body decays (red curve) have a smaller $L_{xy}(D)_B$ than charmed B decays (black curve). From a MC study we verify that the efficiency of the cut $L_{xy}(D)_B > 0.01$ cm is to remove about 75% of background and about 20% of signal events.
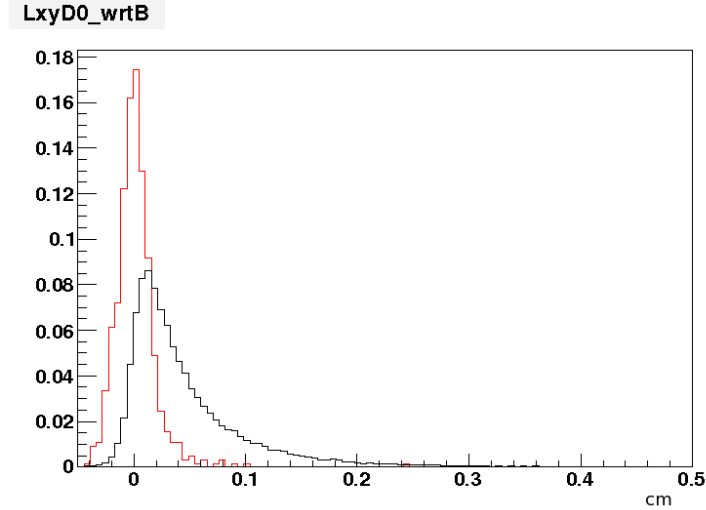
Figure 3.3: Distribution of $L_{xy}(D)_B$ of a three body decay (red curve) and a charmed B decay (black curve).

We also confirm the B_CHARM selection criteria using the off-line variables. This ensure good agreement between Monte Carlo simulation and data, because it removes particular background events from data that are not simulated in the Monte Carlo. They are the so called "volunteers", track pairs in which a track from a real B decay is combined with a fake track, or in which a track from two real Bs are combined.

3. Selection of the suppressed signals $B^- \to D(\to K^+\pi^-)h^-$

After these basic requirements, we build the invariant mass distributions of the B candidates assuming the pion mass hypothesis to the track from B (Fig. 3.4). In this way the $B^- \to D\pi^-$ peak appears in the correct mass (about 5.279 GeV/c^2), while the events from the $B^- \to DK^-$ are expected to cluster in smaller and wider peak, located about 50 MeV/c^2 below the $B^- \to D\pi^-$. The favored and suppressed invariant mass distributions, as obtained after these basic requirements, are shown in the Fig. 3.4. We can clearly see the favored $B^- \to D\pi^-$ peak at the correct mass range, while in the suppressed sample only the combinatorial background is visible, that overcome the suppressed peaks. Therefore a crucial point of this analysis is the background suppression, that can be achieved by an optimized criteria selection.

3.3 Optimized selection

The strategy is to use the favored $B^- \to D\pi^-$ signal as a model for the suppressed one. It has the same topology of the suppressed signal and allows us to not to rely on Monte Carlo simulation, but to directly use data for both signal and background distributions.

For this reason we select the signal region between $\pm 2\sigma$ around the B mass (5.243 GeV/$c^2 \le M(B) \le 5.315$ GeV/c^2), and, as background region, the mass window 5.4 GeV/$c^2 \le M(B) \le 5.8$ GeV/c^2, where only combinatorial background and no physics background appears.

Since we take the distributions from data, in the signal range there are also background events. To remove this events we perform the so called "sideband subtraction". We fit the background window with an exponential function and we evaluate the ratio of its integral in the background and signal regions. This ratio is the scale factor to subtract the background observables from the signal observables.

We maximized the figure of merit $\dfrac{N_S}{a/2 + \sqrt{N_B}}$, where N_S is the number of signal events, N_B the number of background events and a is the desirable sensitivity for the signal observation, in unit of standard deviation σ. We used the value $a = 3$, so the figure of merit assumed the expression: $\dfrac{N_S}{1.5 + \sqrt{N_B}}$. We can demonstrate [65] that
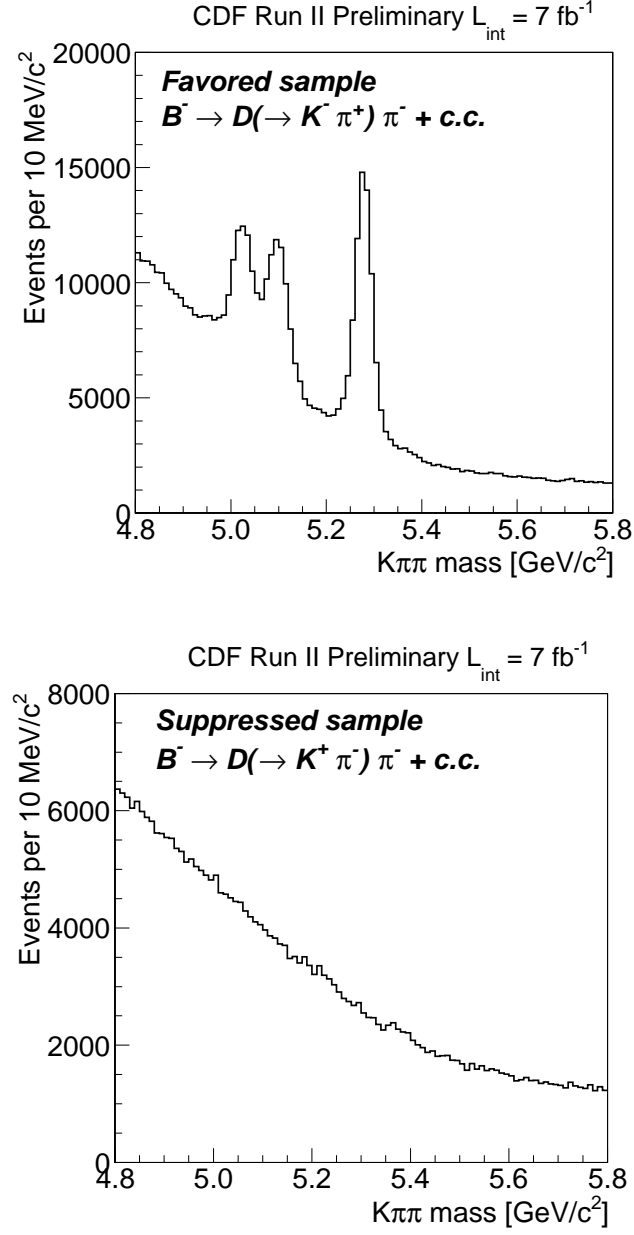
Figure 3.4: Invariant mass distribution of $B^- \to D(\to K^-\pi^+)h^-$ (top) and $B^- \to D(\to K^+\pi^-)h^-$ (bottom) after applying the basic requirements.

through this quantity we can take the minimum amount of necessary data to see the signal with that sensitivity, or to exclude it with the same sensitivity. This quantity is also independent from the knowledge of the cross section of the phenomena, and this make it suitable for any kind of experiment. It has also the advantage to not diverge in the case of low background expectation.

3.3.1 Additional variables used in the selection

We include the following variables in the optimization:

- D Mass: $M(D)$;

- B decay length significance: $\dfrac{L_{xy}(B)}{\sigma_{L_{xy}(B)}}$;

- B impact parameter: $|d_0(B)|$;

- B tridimensional vertex quality: $\chi^2_{3D}(B)$;

- ΔR between the track from B and the D;

- B Isolation (Cone of Radius 1): $Isol_1$;

- B Isolation (Cone of Radius 0.4): $Isol_{0.4}$;

- B pointing angle: $\alpha(B)$;

- D decay angle: $\cos\theta_D^*$;

- Difference of kaoness: $\kappa(K_D) - \kappa(\pi_D)$.

The **isolation** represents an estimator of the fraction of transverse momentum carried by the B meson. B hadrons tend to carry a larger fraction of transverse momentum with respect to lighter hadrons and background [66]. The isolation variable is defined as:

$$Isol_R = \frac{p_T(B)}{p_T(B) + \sum_{i \neq f_B}^{R} p_T(i)}$$

where the sum runs over all tracks inside a cone of radius R around the direction of the B hadron momentum, in the $\eta - \phi$ space, excluding the decay products of the B hadron itself.

The *Pointing Angle α* is defined as the angle between the 3-dimensional momentum of B and the 3-dimensional decay lenght. Signal events will have small pointing angles, while background events will have bigger angles.

The *decay angle θ_D^** is defined as the angle between the D and the flight direction of B in the center of mass frame of the B. We use the $\cos\theta_D^*$ in the optimization procedure (see Fig. 3.7).

The variable $\kappa(i)$, called *kaoness*, is defined as $\kappa(i) = \dfrac{dE/dx_{meas}(i) - dE/dx_{exp}(\pi)}{dE/dx_{exp}(K) - dE/dx_{exp}(\pi)}$, where $i = K, \pi$ from D. More details are given in Sect. 3.3.4.

3.3.2 The separation power

For each distribution we evaluate the *separation power s*.

The conventional way to quote a separation between distributions is to provide an estimate of the distance between the centers of the distributions in units of their standard deviations. However the meaningfulness of this estimation degrades as the distributions deviate from the Gaussian shape. In presence of long tails or strongly asymmetric distributions, the separation estimated with this approach is not of much use.

We use a definition independent on the shape of the distribution, that allows to quickly evaluate the expected resolution on extracting the fractions of events in each category before performing any fit.

The separation between two classes of events is defined [67] as the relative precision with respect to the ideal case $s = \sigma_f^{best}/\sigma_f$, where $\sigma_f^{best} = \left[\sqrt{N/f(1-f)}\right]^{-1}$ and σ_f is given by the Minimum Variance Bound, $\sigma_f = \sqrt{1/N \left[\int \dfrac{(\mathcal{P}_1(x) - \mathcal{P}_2(x))^2}{f\mathcal{P}_1(x) + (1-f)\mathcal{P}_2(x)}dx\right]^{-1}}$, where $\mathcal{P}_{1(2)}(x)$ is the probability of the probability distribution of the event x of class 1(2). The separation power is therefore defined as:

$$s = \sigma_f^{best}/\sigma_f = \sqrt{f(1-f)\int \frac{(\mathcal{P}_1(x) - \mathcal{P}_2(x))^2}{f\mathcal{P}_1(x) + (1-f)\mathcal{P}_2(x)}dx}.$$

This quantity is independent of the sample size, but depends on the true values of fractions as it generally happens for resolutions. It ranges from zero for no separation, i.e. completely overlapping distributions, to one for the maximum achievable separation.

Figs. 3.5, 3.6 and 3.7 show the distributions of the variables listed above with the corresponding separation power for signal, sideband subtracted, and background (in red and black colors respectively).

(a) separation = 0.66

(b) separation = 0.49

(c) separation = 0.45
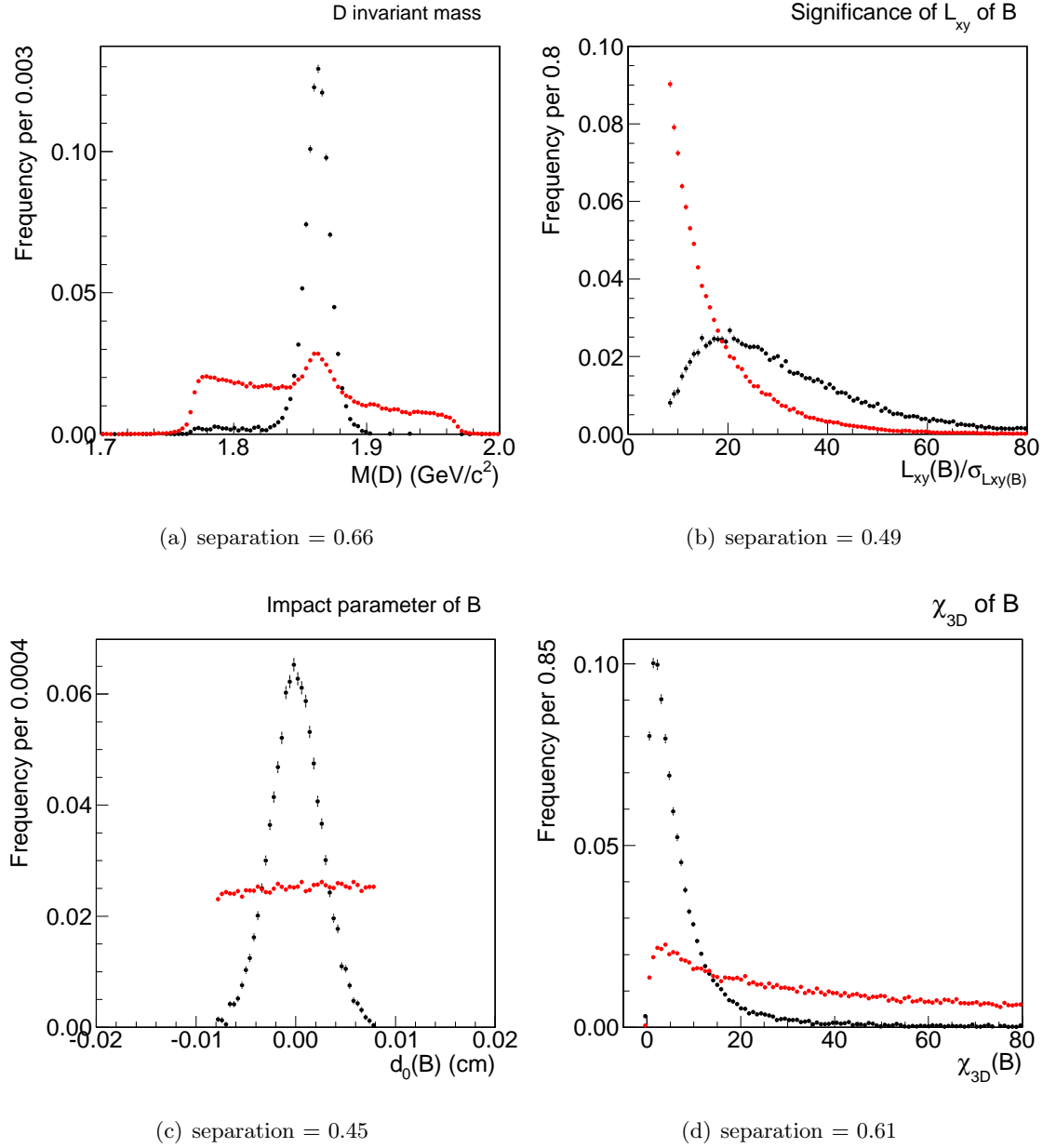
(d) separation = 0.61

Figure 3.5: From left to right we have the distribution of variables D *invariant mass*, $\frac{L_{xy}(B)}{\sigma_{L_{xy}(B)}}$, $d_0(B)$ and $\chi^2_{3D}(B)$ in the signal region (in black) and in the background region (in red).
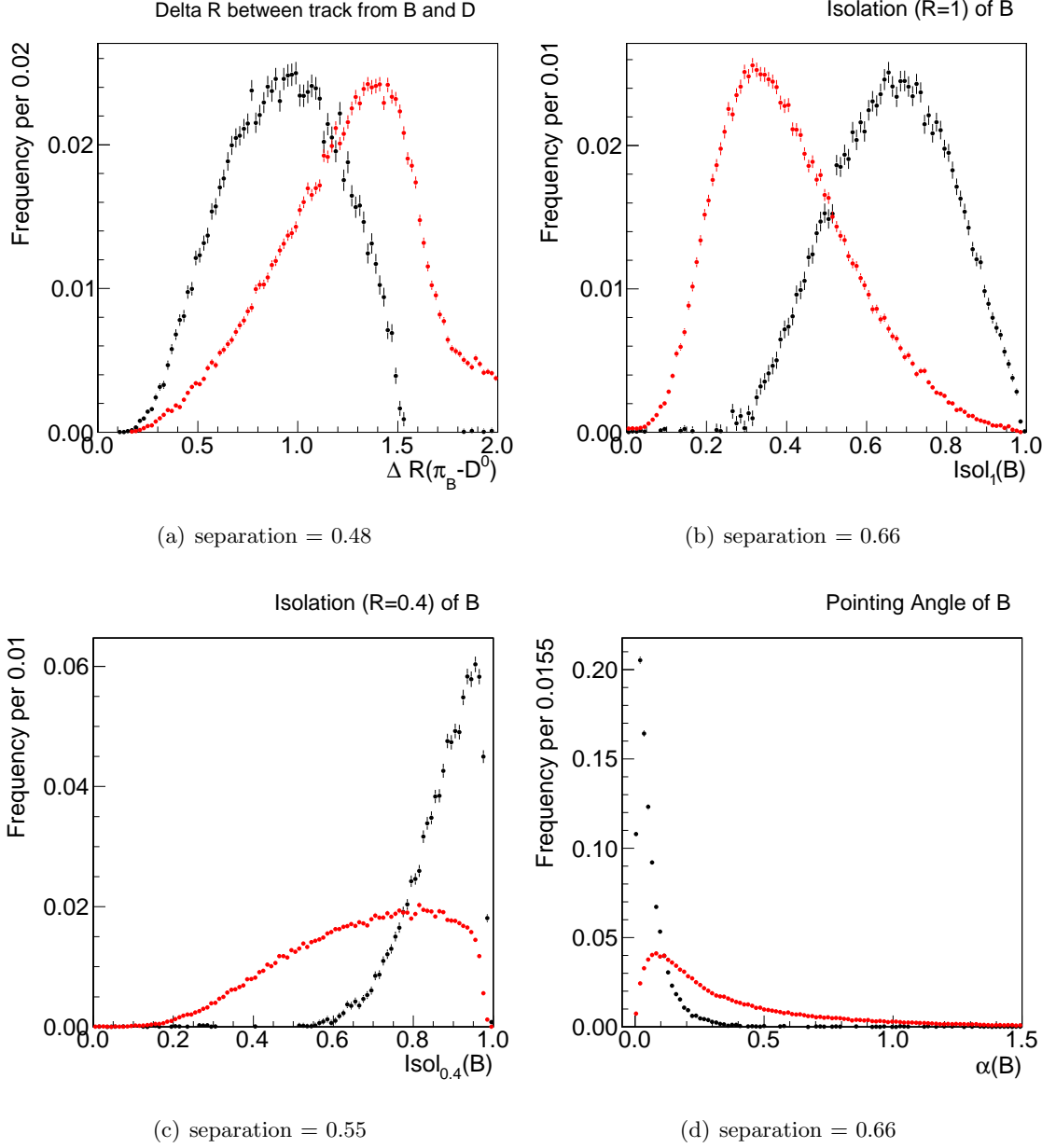
(a) separation = 0.48

(b) separation = 0.66

(c) separation = 0.55

(d) separation = 0.66

Figure 3.6: From left to right we have the distribution of variables ΔR , $Isol_1$, $Isol_{0.4}$, $\alpha(B)$ and $\cos\theta_D^*$ in the signal region (in black) and in the background region (in red).

(a) separation = 0.30

(b) separation = 0.27
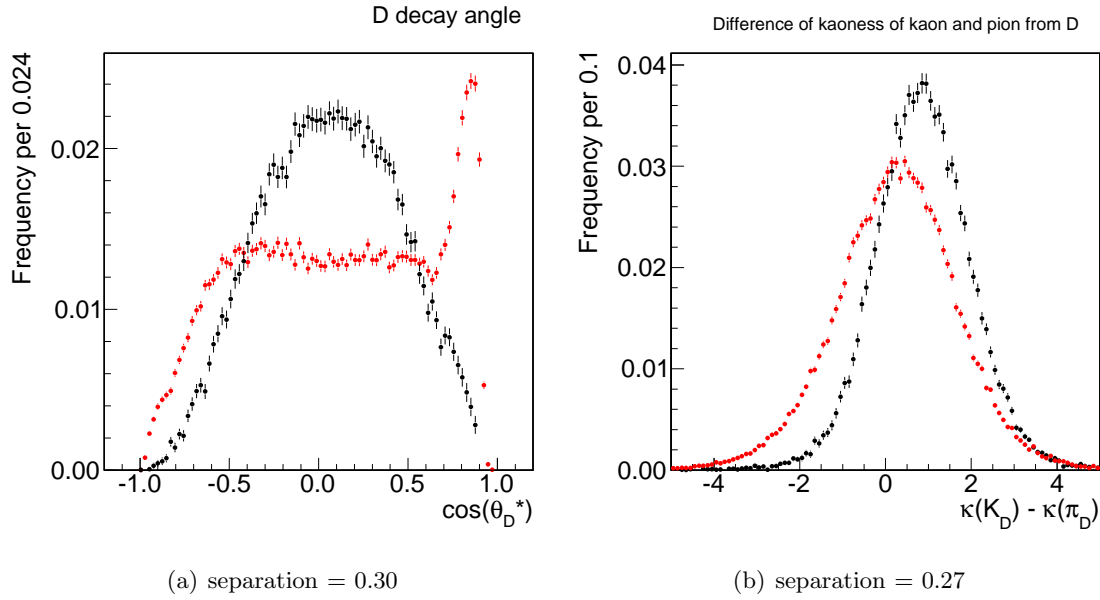
Figure 3.7: From left to right we have the distribution of variables $\cos\theta_D^*$ and $\kappa(K_D) - \kappa(\pi_D)$ in the signal region (in black) and in the background region (in red).

3.3.3 Study on the isolation

We studied the distribution of the isolation at various radii, from $R = 0.3$ to $R = 1.3$, for signal and background events. In Figs. 3.9 and 3.10 the corresponding distributions are reported (in black for the signal and red for the background) with the separation power for each radius.

(a) separation = 0.51

(b) separation = 0.55

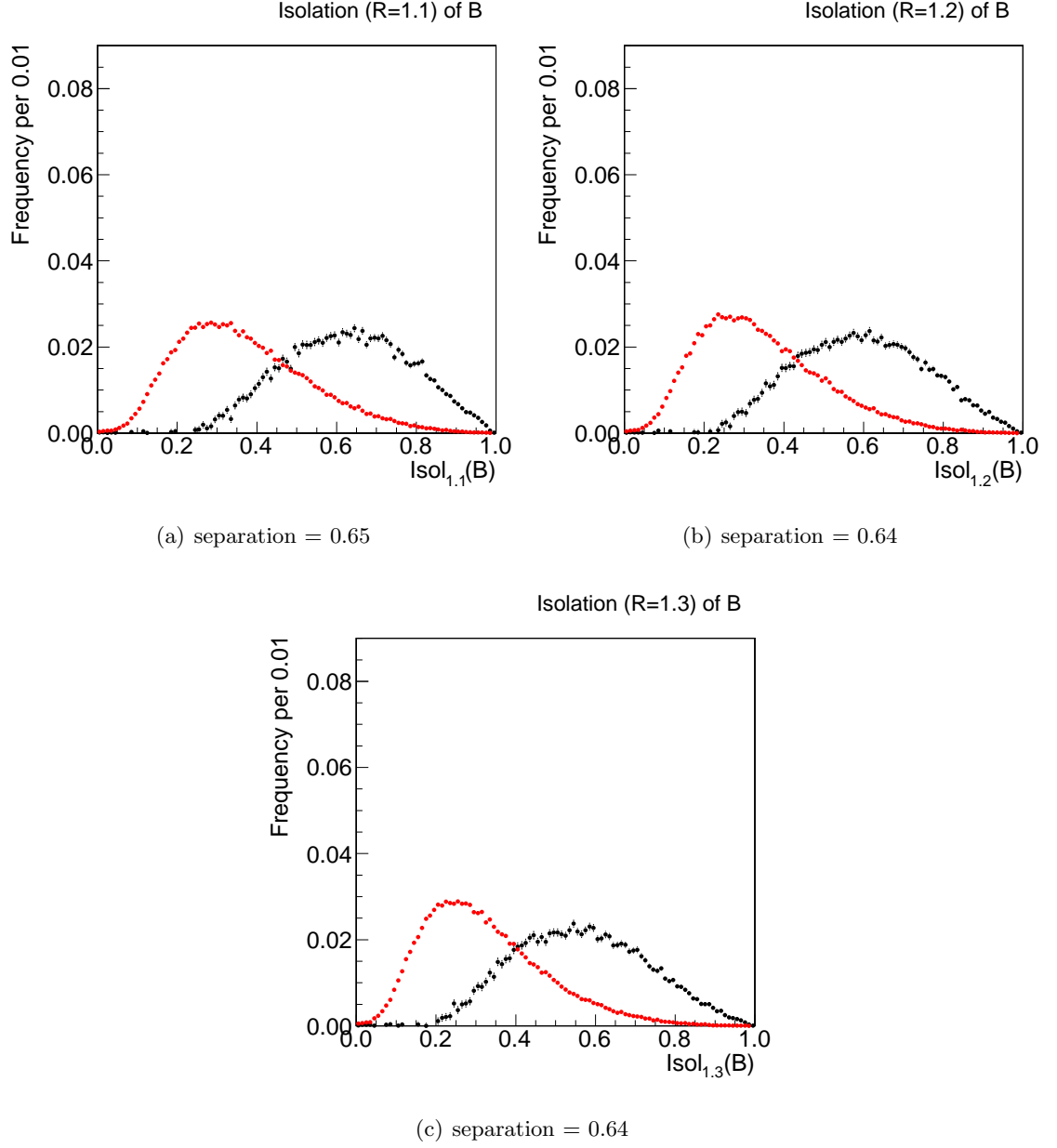(c) separation = 0.60

(d) separation = 0.63
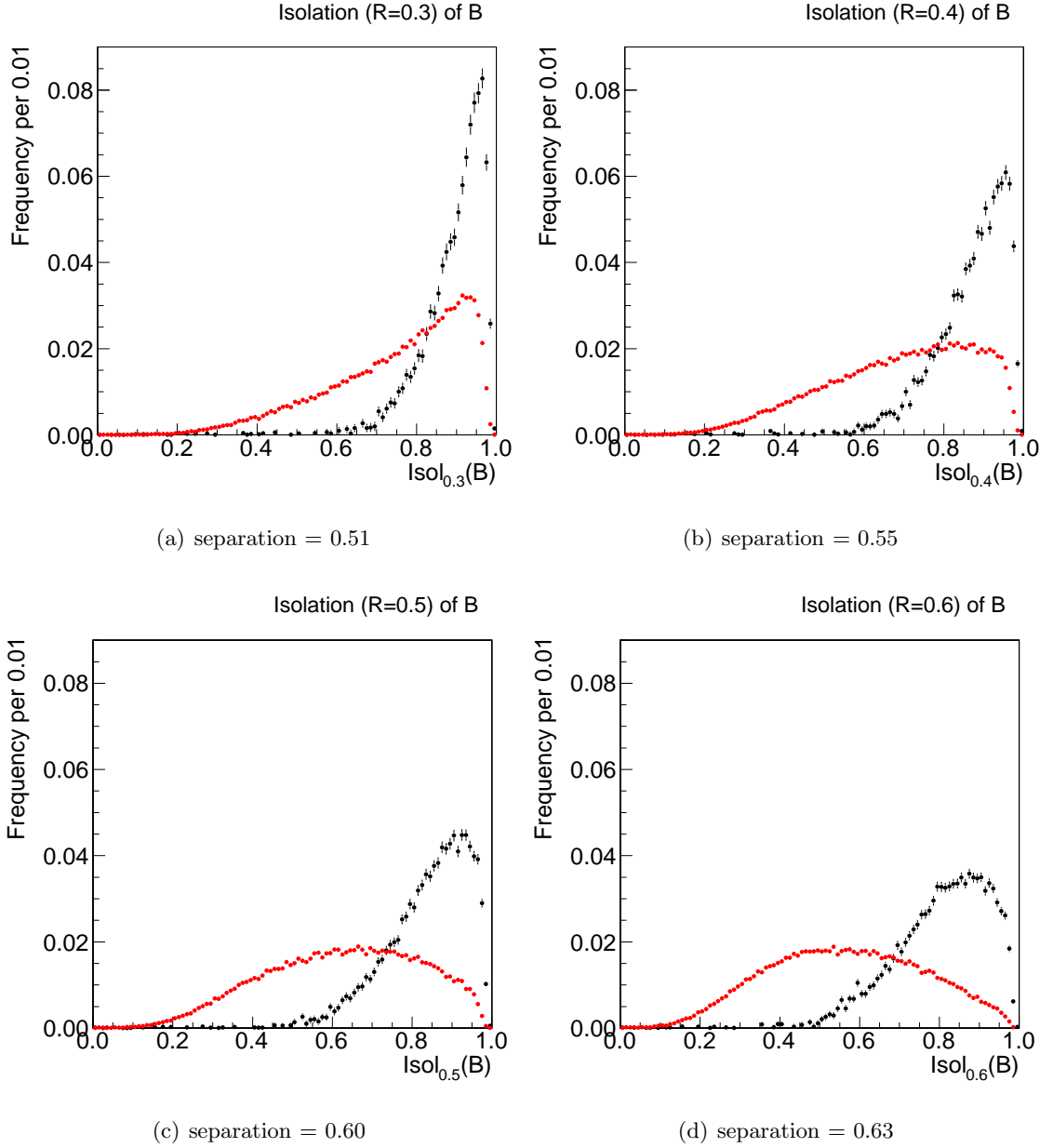
Figure 3.8: From left to right we have the distribution of variables ΔR , $Isol_1$, $Isol_{0.4}$, $\alpha(B)$, $\cos \theta_D^*$ and $\kappa(K_D) - \kappa(\pi_D)$ in the signal region (in black) and in the background region (in red).

(a) separation = 0.65   (b) separation = 0.66

(c) separation = 0.66   (d) separation = 0.66
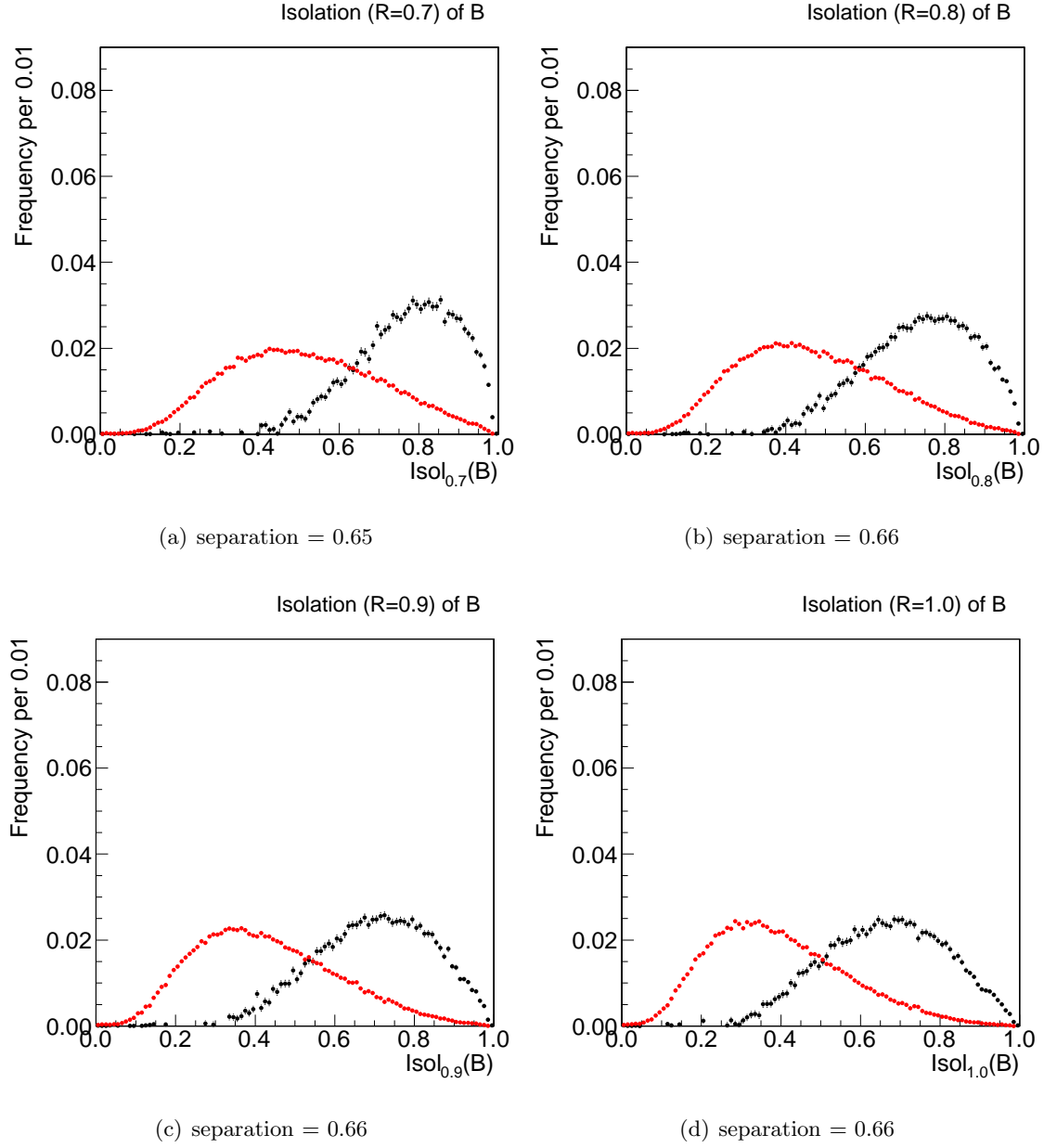
Figure 3.9: From left to right we have the distribution of variables ΔR , $Isol_1$, $Isol_{0.4}$, $\alpha(B)$, $\cos\theta_D^*$ and $\kappa(K_D) - \kappa(\pi_D)$ in the signal region (in black) and in the background region (in red).

A dedicated study has been made to verify a possible improvement of the selection combining several isolation observables.

First of all we combined two single isolations in a statistics called *pseudo Likelihood-ratio, pRL*. The pLR combines the information of the interesting variables with the product of the Likelihood ratio LR of the single variables, and does not consider the correlations among them. pRL and LR coincide in case of uncorrelated variables.

(a) separation = 0.65

(b) separation = 0.64

(c) separation = 0.64

Figure 3.10: From left to right we have the distribution of variables ΔR , $Isol_1$, $Isol_{0.4}$, $\alpha(B)$, $\cos\theta_D^*$ and $\kappa(K_D) - \kappa(\pi_D)$ in the signal region (in black) and in the background region (in red).

If we have a multi-dimensional event $\vec{x} = (x_1, \ldots, x_n)$, where the coordinates are the values of the variables characterizing the event, and supposing to have two hypothesis H_0 and H_1, for which the event $\vec{x}$ is distributed with a probability density function, pdf, respectively $f_0(\vec{x})$ and $f_1(\vec{x})$, the LR is defined as the ratio between the two pdfs: $LR = \frac{f_0(\vec{x})}{f_1(\vec{x})}$. In the hypothesis of independent variables we can write

the pLR as the product of the pdfs:

$$pLR = \prod_i \frac{f_0(\vec{x})}{f_1(\vec{x})},$$

where the index i runs over the variables x_i characterizing the events.

In our case H_0 corresponds to the "signal" hypothesis and H_1 corresponds to the "background" hypothesis. The expression is better written using the logarithm of the pLR, for which only sums are involved, rather than products.

$$pLR = log \prod_i \frac{pdf_{signal}(x_i)}{pdf_{background}(x_i)} = \sum_i log(pdf_{signal}(x_i)) - \sum_i log(pdf_{background}(x_i)) \ .$$

$$(3.1)$$

The LR allows to compare the likelihood of an event to be identified as signal with the likelihood of being identified as background. Higher values of LR correspond to higher probabilities to have a signal event, while lower values of LR correspond to higher probabilities to have a background event.

The isolation variables are strongly correlated to each other, but those based on very different radii can potentially provide independent information. We select for example the isolations at radii $R = 0.4$ and $R = 1$. To make things easier, we perform an initial change of variables to reduce their correlation:

$$\begin{cases} Isol_1 & \to & 1/Isol_1 \\ Isol_{0.4} & \to & 1/Isol_{0.4} - 1/Isol_1 \end{cases}$$

The initial variables have a correlation factor[4] of 0.69 for the signal and 0.57 for the background, while the new variables have a correlation of 0.47 and 0.15 respectively. We parameterize their distributions with some polynomial functions, and we evaluate the pLR and their separations. Starting from a separation of $1/Isol_{0.4}$ of 0.54 and $1/Isol_{0.4} - 1/Isol_1$ of 0.56, the separation of the pLR is 0.62. It has therefore a better separation than the two single variables. But unfortunately it has a worse separation than the $Isol_1$ alone (0.65). This suggests that there is no easy way to obtain a large improvement in separation by combining different isolation variables, at least not without paying close attention to model their correlations.

---

[4]The correlation factor is the ratio between the covariance of two variables and the product of their standard deviations.

Another study has been done, trying to include the information of all the isolations in a multivariate technique. In particular we used the MultiLayer Perceptrons network (MLP). The *perceptrons* are the simple *neurons* of the network, which compute a single output from multiple real-valued inputs by forming a linear combination of the *weights* and then possibly putting the output through some nonlinear activation function. A typical MLP network consists of a set of source nodes forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes. The input signal propagates through the network layer-by-layer. MLP networks are typically used in *supervised learning* problems. This means that there is a training set of input-output pairs and the network must learn to model the dependency between them. The training here means adapting all the weights and biases to their optimal values for the given pairs. The supervised learning problem of the MLP can be solved with the *back-propagation algorithm*. The algorithm consists of two steps. In the *forward pass*, the predicted outputs corresponding to the given inputs are evaluated. In the *backward pass*, partial derivatives of the cost function with respect to the different parameters are propagated back through the network.

The final result [68] shows no significant improvements if considering the MLP output rather then the single isolations: the separation is still close to the value of the isolation at $R = 1$.

Finally we empirically tried to use two of them in a rectangular optimization of the cuts. When using only one of them, the $Isol_1$, we obtain a signal over background ratio in the suppressed sample of about 0.22. Using $Isol_1$ and $Isol_{0.4}$ we improve that value to 0.25. Using both of them we reduce the background events of about 25%, while the signal events of about 12%. For these reasons we decide to use both $Isol_1$ and $Isol_{0.4}$ in the selection.

3.3.4   Study on the kaoness

It is known that the separation of residual [5] dE/dx information of kaons and pions with momenta greater than 2 GeV/c is about 1.5 σ [36]. The distributions are

---

[5]The residual is defined as the difference of the measured dE/dx of the track and the expected value, $dE/dx_{meas} - dE/dx_{exp}$
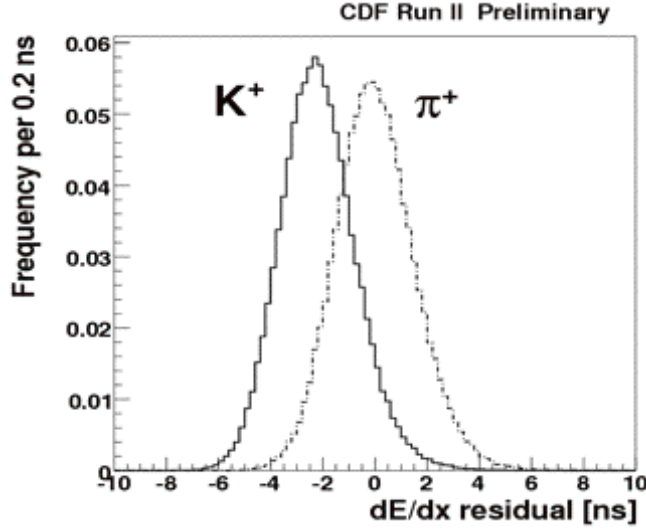
reported in Fig. 3.11.

Figure 3.11: Residual of dE/dx measurement for kaons and pions with momentum greater than 2 GeV/c. The separation is about 1.5 σ.

To combine the information of the energy loss in the kaoness variable (see Sect. 3.3.1) has several advantages. In particular it is at the first order momentum independent and it has the feature to have the mean value equal to 0 in case of pions and equal to 1 in case of kaons. In Fig. 3.12 we can see the distributions of the kaoness of kaons (pions) from D calculated for signal and background events. The separation is about 0.32 (0.23). The distributions are taken using tracks with momentum greater than 2 GeV/c. When using also momentum lower than 2 GeV/c we obtain the distributions in Fig. 3.13 where the separations are respectively 0.27 and 0.20. The use of tracks with momentum lower than 2 GeV/c reduce the performances of the separation, because the dE/dx calibrations are not optimized for those tracks. However the addition of those tracks still improve the overall power of the sample.

Another possibility that we have is to use the difference of the two kaoness $\Delta\kappa = \kappa(K_D) - \kappa(\pi_D)$. This variable is shown in Fig. 3.14 and has a separation of 0.30 to discriminate signal from background.

To choose whether it is better to use the two kaoness or the combination in the subtraction, we do the optimization in both cases. Using the two singles, we find
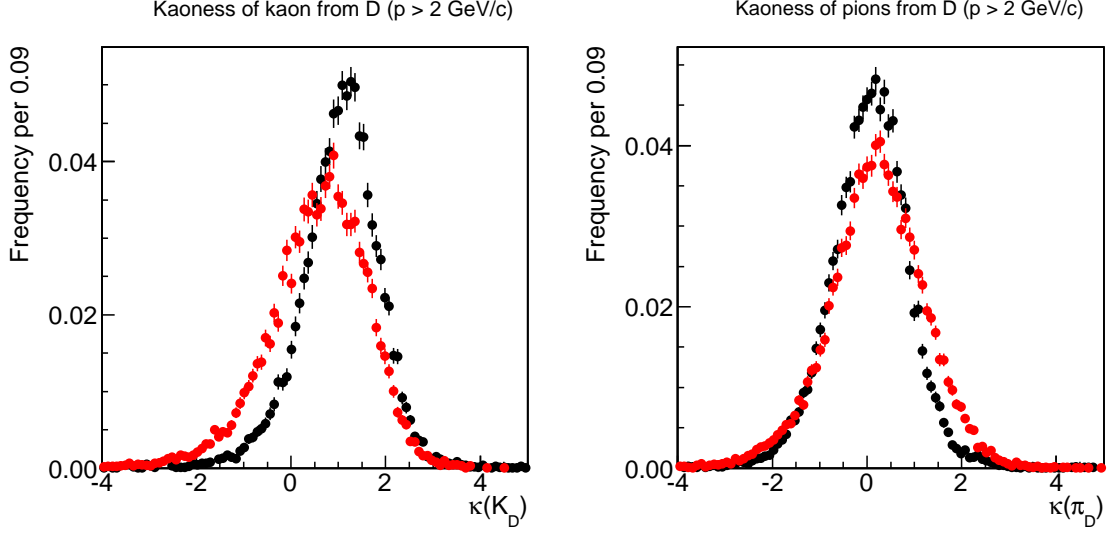
Figure 3.12: Kaoness of kaons (left) and pions (right), with momentum greater than 2 GeV/c, from D calculated for signal (black) and background (red) events. The separation is 0.32 and 0.23, respectively.
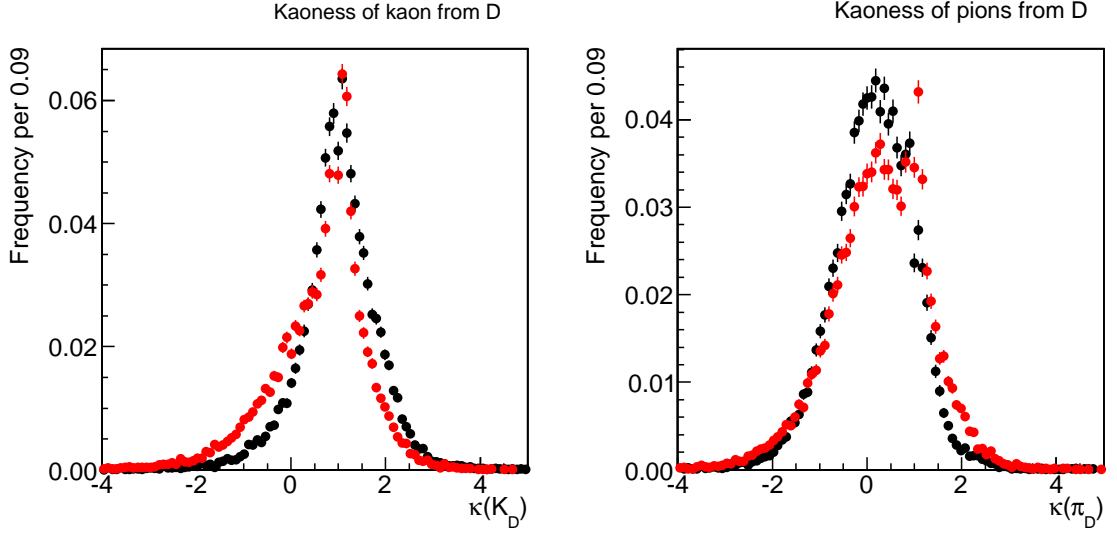

Figure 3.13: Kaoness of kaons (left) and pions (right) from D calculated for signal (black) and background (red) events. The separation is 0.27 and 0.20, respectively.

a signal over background ratio of about 0.30, while using the subtraction we find a ratio of about 0.40. Moreover the background is reduced of about 15%, while the signal is increased of about 12%. For these reasons we use the difference of the kaoness in the final selection.
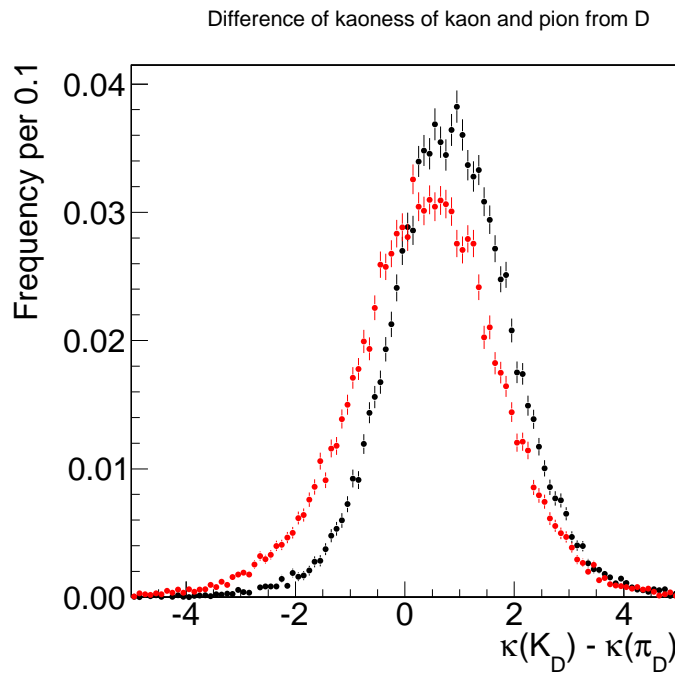
Figure 3.14: Difference of kaoness of kaon and pion from D. In black the signal events and in red the background events. The separation is 0.30.

3.3.5 Optimized criteria

The final selection criteria that maximize the figure of merit is the following:

- D mass: $M(D) \leq 1.8645 + 1.5 \times 0.01$ and $M(D) \leq 1.8645 - 1.5 \times 0.01$;

- B decay length significance: $\dfrac{L_{xy}(B)}{\sigma_{L_{xy}(B)}} \geq 12$;

- B impact parameter: $|d_0(B)| \leq 0.005$ cm;

- B tridimensional vertex quality: $\chi_{3D} \leq 13$;

- $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the track from B and the D: $\Delta R \leq 1.5$;

- B Isolation (Cone 1): $Isol_1 \geq 0.4$;

- B Isolation (Cone 0.4): $Isol_{0.4} \geq 0.7$;

- B pointing angle: $\alpha(B) \leq 0.15$;

- D decay angle: $|\cos(\theta_D^*)| \leq 0.6$;

- Difference $\Delta(\kappa(K_D) - \kappa(\pi_D)) \geq -1$.

The resulting invariant mass distributions of $B^- \to D(\to K^-\pi^+)h^-$ (fav) and $B^- \to D(\to K^+\pi^-)h^-$ (sup), with pion mass assignment to the track from B, were reported in Fig. 3.15.

As in the distribution before the optimized selection (Fig. 3.4), the favored $B^- \to D\pi^-$ signal can be clearly seen at the correct B mass value. But moreover now the combinatorial background (right sideband) is significantly reduced and in the left sideband, around the value of 5.23 GeV/c^2, an hint of $B^- \to DK^-$ signal is visible. Going on lower values of the mass we expect to see decays coming from the so called "partially reconstructed" decays of the type $B^- \to D^{0*}X^-$, where the $D^{0*} \to D^0\pi^0/\gamma$ and the neutral particle is lost in the reconstruction.

In the suppressed sample is now visible a signal structure in the corresponding signal region. For lower values of the mass we expect again partially reconstructed events, even if they don't show the same structure as of the favored sample. The combinatorial level is now about 1/3 of the signal level.

CDF Run II Preliminary L$_{int}$ = 7 fb^{-1}

Favored sample
$B^- \to D(\to K^-\pi^+)\,\pi^-$ **+ c.c.**

Events per 10 MeV/c^2

K$\pi\pi$ mass [GeV/c^2]

CDF Run II Preliminary L$_{int}$ = 7 fb^{-1}

Suppressed sample
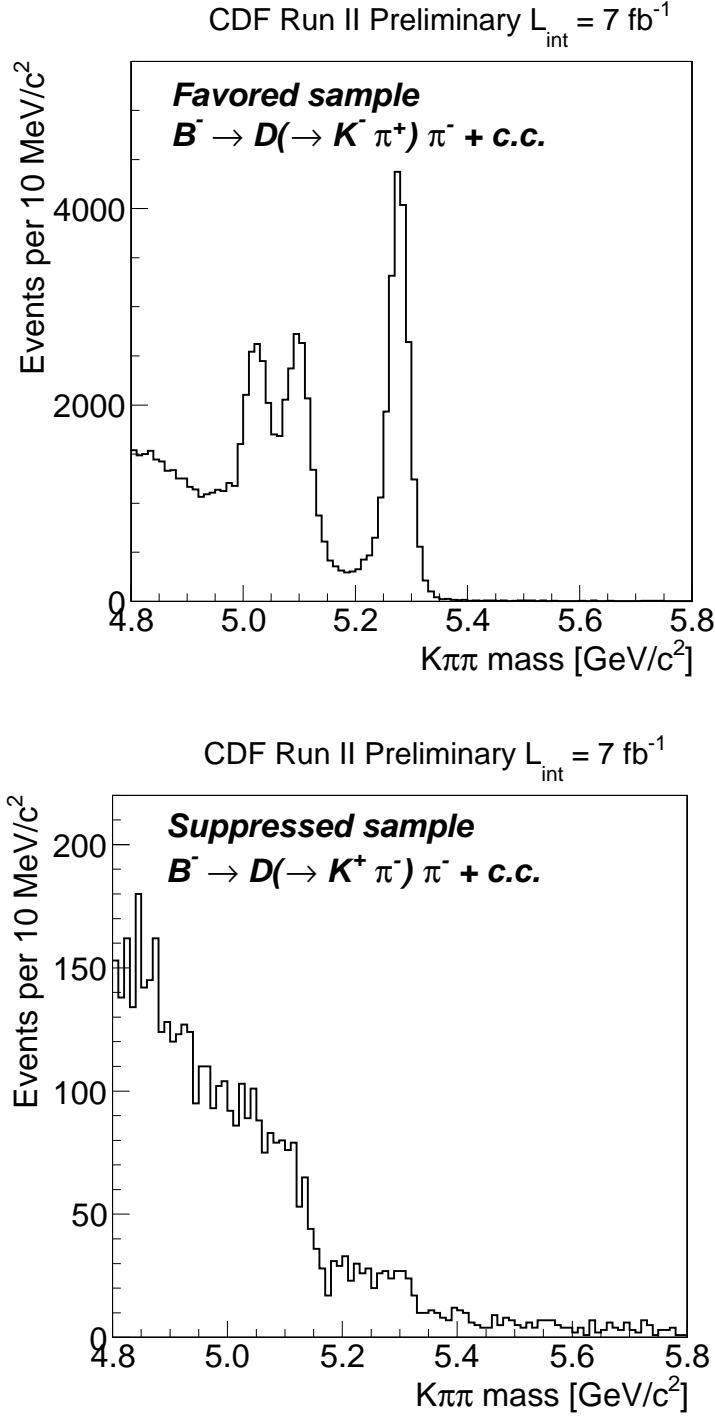$B^- \to D(\to K^+\pi^-)\,\pi^-$ **+ c.c.**

Events per 10 MeV/c^2

K$\pi\pi$ mass [GeV/c^2]

Figure 3.15: Invariant mass distribution of $B^- \to D(\to K^-\pi^+)\pi^-$ (top) and $B^- \to D(\to K^+\pi^-)\pi^-$ (bottom), with a nominal pion assignment to the charged track from the B meson decay, after optimized selection.

To know the exact composition of the physical backgrounds in the two samples we need to look at the Monte Carlo (MC) simulation, that models both the pro-

duction process of a B decay and its interaction with the detector material. The same selection criteria of the analysis is applied to MC samples and the relative contributions to the yields are extracted from a fit.

Details about these procedures are given in the next chapter.

Simultaneous fit of $B^- \to DK^-$ and $B^- \to D\pi^-$ modes

In this chapter we will describe which are the backgrounds of our favored and suppressed samples and the implementation of an extended maximum Likelihood fit that, combining the mass and particle identification information, provides the statistical separation of the $B^- \to Dh^-$ modes from the backgrounds. The fit is simultaneously applied to both samples.

4.1 Monte Carlo simulation

To find which are the background contributions to our sample, we simulate events though Monte Carlo techniques, simulating the generation, the particle decay and the interaction of the particle in the detector material (see Sect. 2.10).

For the signal we generate Monte Carlo sample of $B^- \to Dh^-$ decays including also the effect of QED radiation in the final state with the loss of energy from soft photon emission from charged final state particles. Since the precision on measurement involving B and D mesons is increased in the recent years, Monte Carlo simulation has to account also for the effects for emission of soft photons in processes with charged particles [69], which are of the order of per-cent.

This effect is well modeled in Monte Carlo through PHOTOS [70, 71], an algo-

rithm that includes leading and next to leading order calculations of the QED effects of the soft-photon radiation. This allow to have a precise model of the mass shape of the decays, at a level of 0.01%.

Since we observe a discrepancy between the $p_T(B)$ distributions of the simulated B candidates and of B candidates reconstructed on real data (Fig. 4.1), we re-weighted the distributions with a weight function given by the ratio of $p_T(B)$ from data and MC.

To do this we selected the signal region in a window of $\pm 2\sigma$ around the B peak, sideband subtracted. For the sideband subtraction we used the same strategy described in Sect. 3.3. The ratio of $p_T(B)$ data/MC is fitted with the complementary "error" function[1] plus a first degree polynomial (see Fig. 4.2) and it is used to re-weight the $p_T(B)$ distribution of the simulated events accordingly. The agreement between MC and data after the reweighting is now satisfactory (Figs. 4.3-4.8). In all the Figs. we show data as points with error bars and MC as lines.
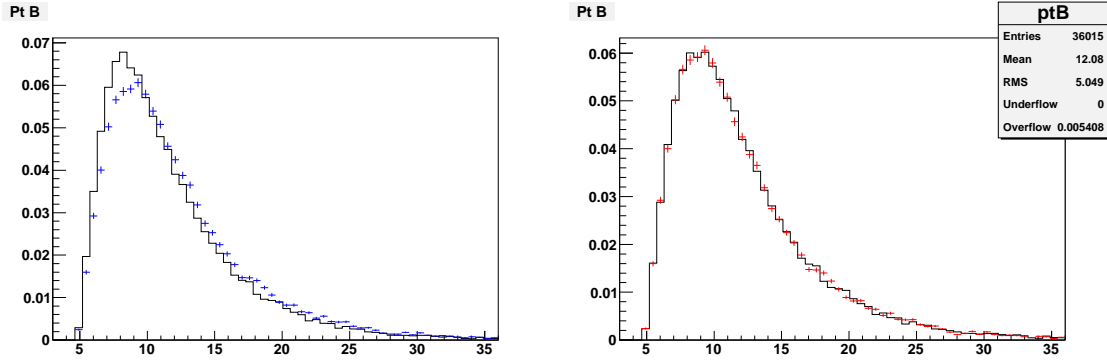

Figure 4.1: Data (points) - simulation (line) comparison of the $p_T(B)$ distribution. Left plot: $p_T(B)$ distributions before the re-weighting. Right plot: $p_T(B)$ distributions after the re-weighting.

---

[1]The complementary error function is defined as $erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt$.
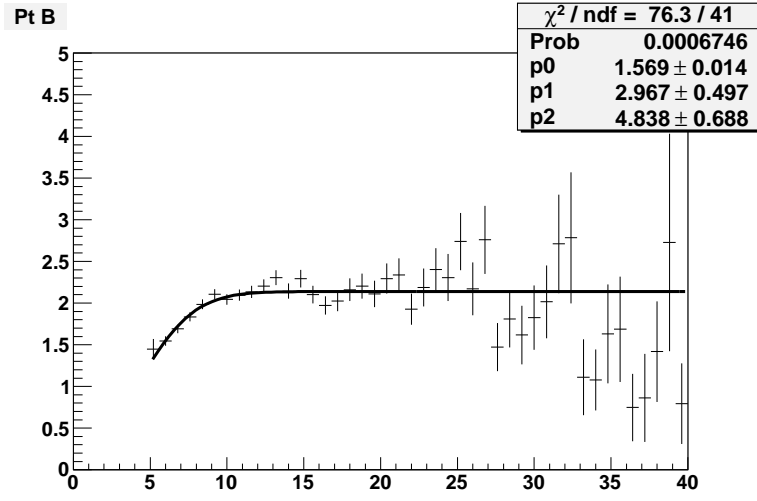
Figure 4.2: Distribution of the ratio data(points)/MC(line) fitted with a complementary error function plus a first degree polynomial.
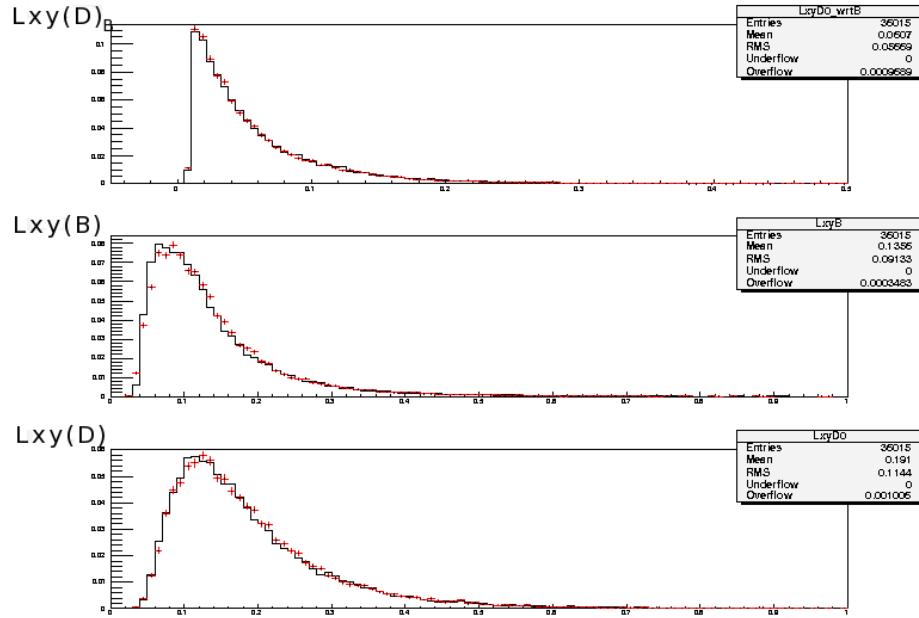

Figure 4.3: Top: D decay lenght measured with respect to the B deacy vertex. Middle: B decay lenght. Bottom: D decay lenght. The line is from the MC, points are data.
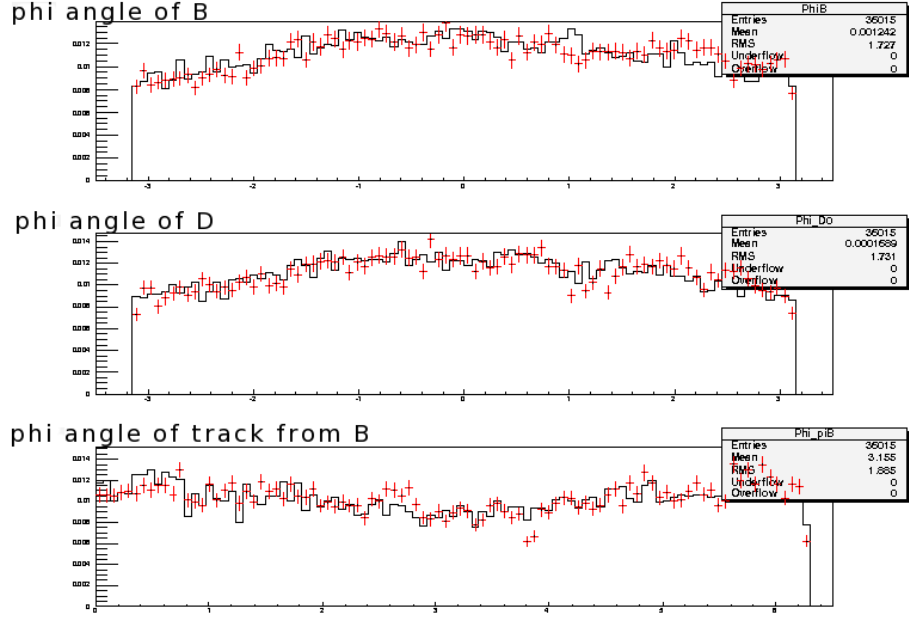
Figure 4.4: Top: ϕ angle of B. Middle: ϕ angle of D. Bottom: ϕ angle of the pion from B. The line is from the MC, points are data.
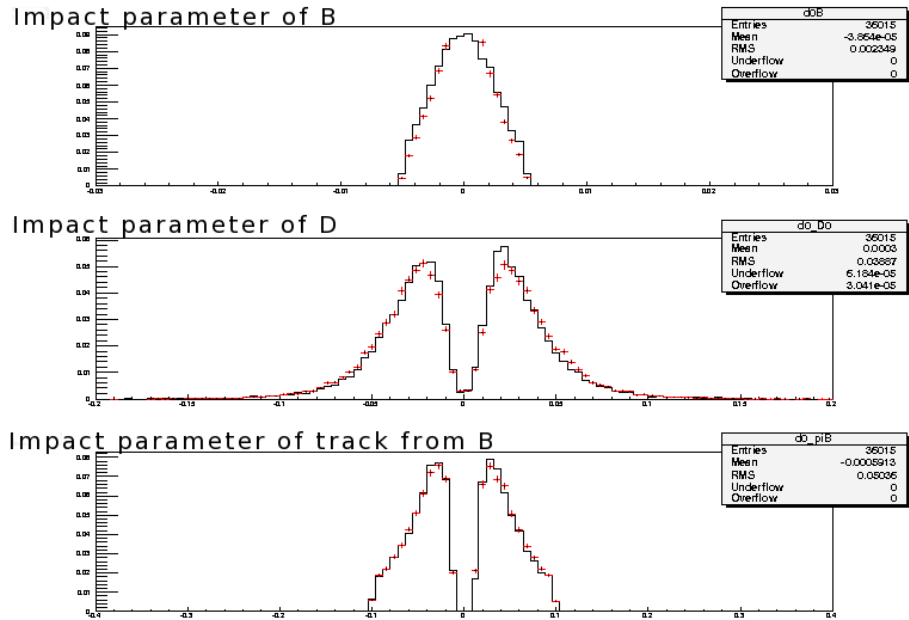

Figure 4.5: Top: B impact parameter. Middle: D impact parameter. Bottom: Impact parameter of the pion from B. The line is from the MC, points are data.
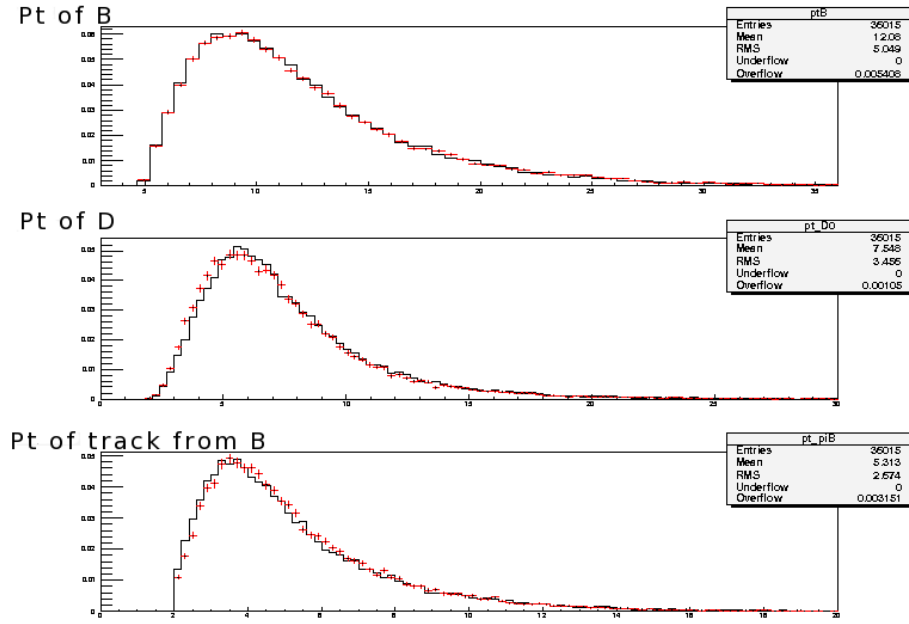
Figure 4.6: Top: $p_T(B)$. Middle: $p_T(D)$. Bottom: $p_T(\text{pion from } B)$. The line is from the MC, points are data.
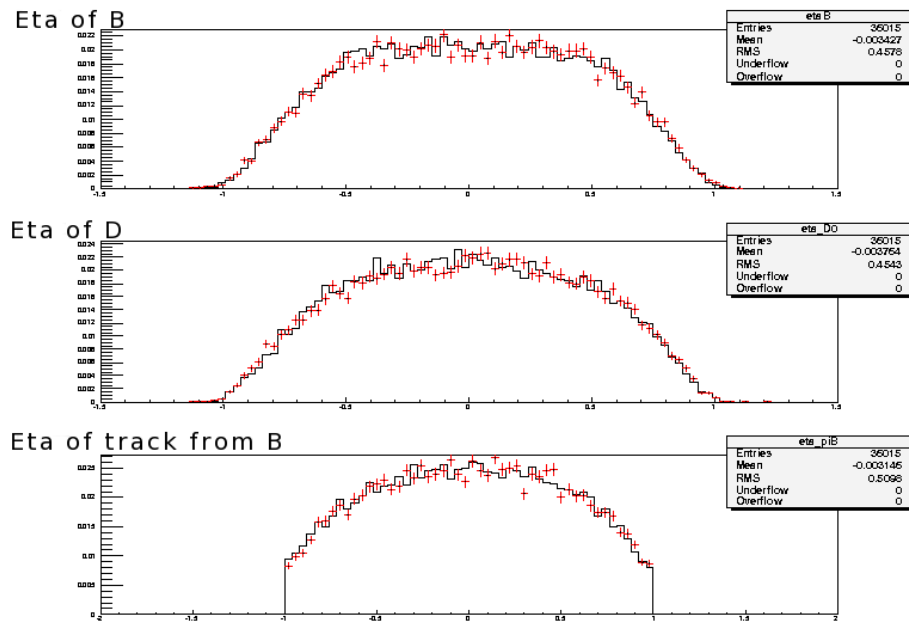

Figure 4.7: Top: η angle of B. Middle: η angle of D. Bottom: η angle of the pion from B. The line is from the MC, points are data.
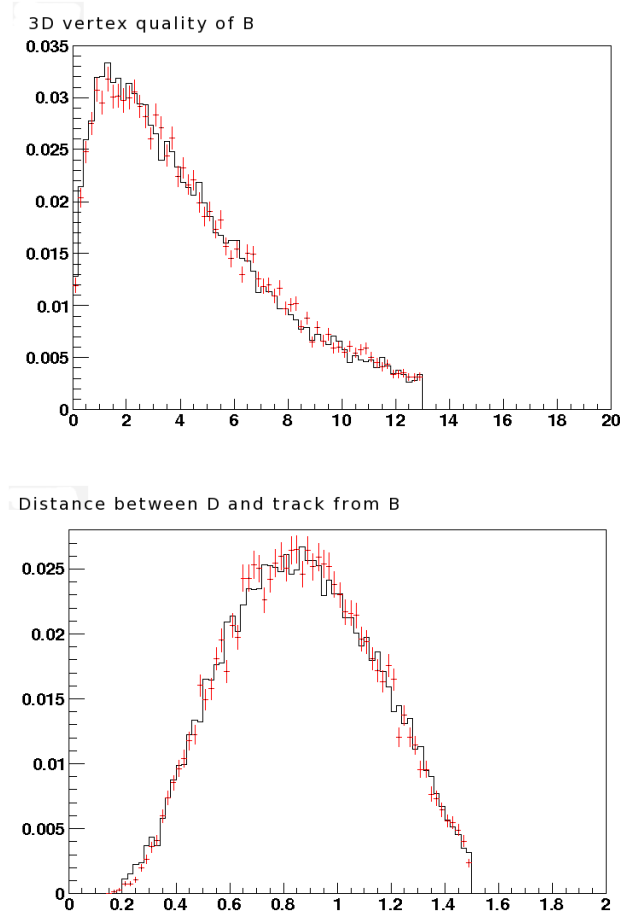
Figure 4.8: Left: tridimensional χ^2. Right: $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the D the track from B. The line is from the MC, points are data.

4.1.1 Background study for the favored sample

For the favored sample we use a generic sample of 6000000 B^- decays, with the request that $D^0 \to K^- \pi^+$. The events are reconstructed with the pion assignment to the track from B (Fig. 4.9).
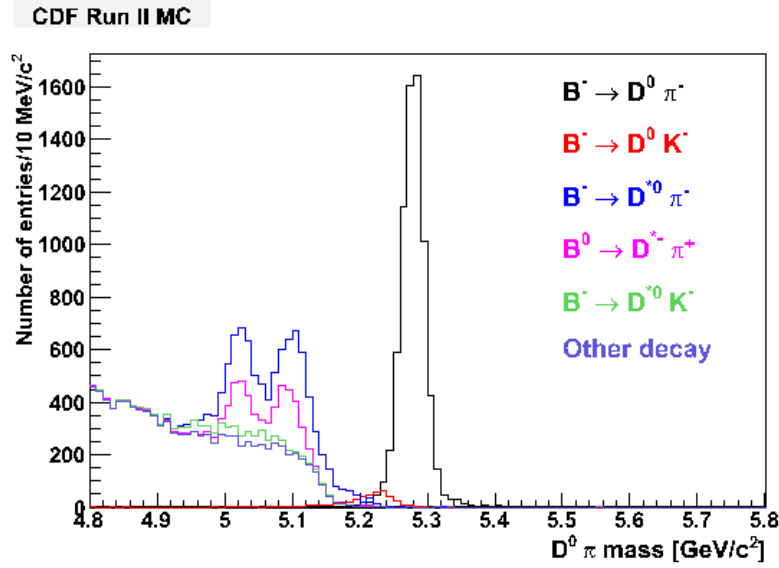

Figure 4.9: Monte Carlo simulated events of generic favored B^- decays.

The global distribution is very similar to the one obtained on data (Fig. 3.15), and we can see the contribution coming from the $B^- \to D\pi^-$ and $B^- \to DK^-$ signals, as well as the partially reconstructed decays on the left.

We note that if we choose as lower limit 5.17 GeV/c^2 we can take into account only the $B^- \to D^*\pi^-$, without considering the other partially reconstructed decays. As higher limit we choose the value 6.5 GeV/c^2 to include more combinatorial background events in the right tail of the mass distribution. This is necessary because, after the optimized selection, the level of the combinatorial background is very low and we need to have a major lever to fit in the right way the slope and its normalization.

4.2 Background study for the suppressed sample

There are several physics backgrounds entering in the suppressed reconstruction. We will consider each category separately.

4.2.1 $B^- \to D(\to X)\pi^-$

The favored mode itself is an important source of background for the suppressed mode. We checked from MC that favored events in which the D radiate a photon in the final state, if reconstructed as suppressed, can escape the veto on the swapped mass (see Sect. 3.2.1). We estimate the fraction of these events, with respect to the favored $B^- \to D\pi^-$, to be about $1.4 \cdot 10^{-3}$. Since the expected suppressed $B^- \to D\pi^-$ branching fraction with respect to the favored is about $3.5 \cdot 10^{-3}$ the contribution of this background is significant.

We find that many other decays behave as background for the suppresed, of the same type of $B^- \to D\pi^-$, with $D \to \pi^-\mu^+\nu_\mu$, $D \to K^-\pi^+\pi^0$, $D \to \rho^-\pi^+$, $D^0 \to \rho^+\pi^-$.

We decide to treat all those decays together as a single background component, generating a MC sample of $B^- \to D\pi^-$ with $D \to X$, where X are mode other than $K^+\pi^-$, with relative branching ratios (BR) as given by official tables of the B-group [72]. Considering all backgrounds together has the advantage to increase the resolution in the final fit, since we fit only the global normalization, while all contributions are each other fixed to the relative BRs.

We evaluate the background fraction with respect to the favored $B^- \to D\pi^-$ to be about $5.1 \cdot 10^{-3}$. Expecting to have about 19000 favored $B^- \to D\pi^-$ events, this fraction corresponds to about 100 events in the mass fit window of the suppressed sample.

4.2.2 $B^- \to D(\to X)K^-$

As in the previous case, we generate a MC sample of $B^- \to D(\to X)K^-$ with relative D BRs as described in [72]. Also in this case X are mode other than $K^+\pi^-$. We evaluate the fraction of this background with respect to the favored

$B^- \to DK^-$ mode to be about $2.3 \cdot 10^{-3}$, which corresponds to about 5 events in the suppressed mass fit window.

4.2.3 $B^- \to D_{CP}^{0(*)} h^-$ with $D^0 \to K^+ K^-, \pi^+ \pi^-$

We verified using MC samples that none of $B^- \to D_{CP}^{0(*)} h^-$ modes, where $D^{0(*)}$ can be D^0 or D^{0*}, reconstructed as suppressed events, pass the selection of the suppressed sample. In case of $D \to KK$ events, being reconstructed as $K\pi$, they go outside the selected D mass window, as can be seen in Fig. 4.10. For the $D \to \pi\pi$ modes the only possibly contribution comes from the $B^- \to D(\to \pi^+\pi^-)K^-$, but, after all the cuts and normalizing for the BR, it is negligible.
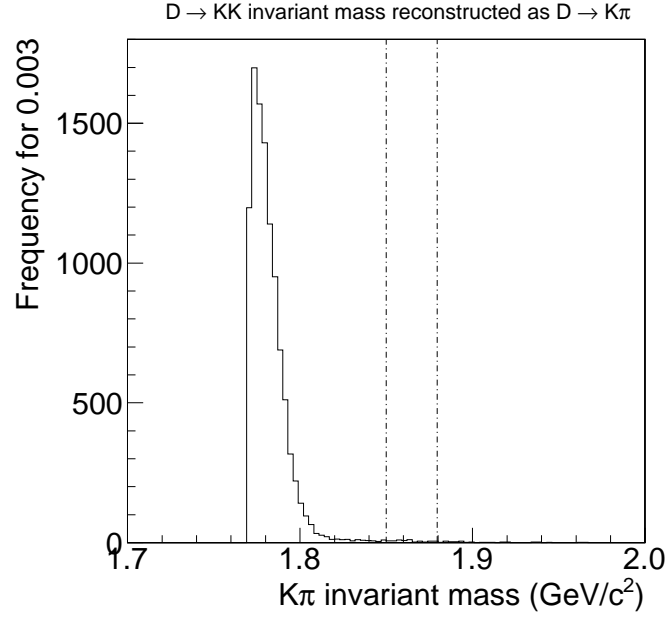

Figure 4.10: Invariant mass of simulated $D \to KK$ events from B^-, reconstructed in the $D \to K\pi$ mass hypothesis. The lines delimit the region selected in our analysis. The $D \to KK$ events are outside that region.

4.2.4 Backgrounds from B^- to three-body decays

The only two 3-body decays that can affect our measurement are $B^- \to K^+\pi^-K^-$, peaking in the $B^- \to DK^-$ mass region, and $B^- \to K^+\pi^-\pi^-$, peaking in the $B^- \to D\pi^-$ mass region.

The first is strongly suppressed by the cut on $L_{xy}(D)_B$. The second one has a fraction of events with respect to favored $B^- \to D\pi^-$ events of about $5.5 \cdot 10^{-4}$, which corresponds to about 11 events in the mass fit window and it will be included as an additional component in the fit.

4.2.5 $B^0 \to D_0^{*-}\ell^+\nu_\ell$

In the decay $B^0 \to D_0^{*-}\ell^+\nu_\ell$, the $D_0^{*-} \to D^0\pi^-$, with $D^0 \to K^+\pi^-$, ignoring the lepton, the final state is $[K^+\pi^-]\pi^-$, mimicking the suppressed mode. We generate MC samples of $B^0 \to D_0^{*-}\ell^+\nu_\ell$ and evaluate the fraction with respect to the favored $B^- \to D\pi^-$ as about $2.4 \cdot 10^{-4}$, corresponding to about 5 events in the suppressed mass window.

4.2.6 Other decays

We generate MC samples of $B_s \to D_s^-\pi^+$, with $D_s^- \to X$, $B^0 \to D^-\pi^+$, with $D^- \to X$, $B^- \to D^0\rho^-$ with $D^0 \to K^-\pi^+$, $B^- \to D^0\mu^-\nu_\mu$ with $D^0 \to K^-\pi^+$, $B^- \to D^0e^-\nu_e$ with $D^0 \to K^-\pi^+$, $B^- \to D^{*0}K^-$ with $D^0 \to K^-\pi^+$, $B^- \to D^{*0}\mu^-\nu_\mu$ with $D^0 \to K^-\pi^+$, $B^- \to K^+\pi^-K^-$ and we find that none of them gives a significant contribution to the suppressed sample.

In conclusion, the four backgrounds that we are going to explicitly consider for the suppressed sample are: $B^- \to D(\to X)\pi^-$, $B^- \to D(\to X)K^-$, $B^- \to K^-\pi^+\pi^-$ and $B^0 \to D_0^{*-}\ell^+\nu_\ell$.

They will be included in a maximum likelihood fit with a Gaussian constraint on their fractions. The mean value of the constraint is the fraction value, while as sigma is taken the sum in quadrature of the error propagated from the BRs coming from PDG, and the relative error coming from the MC statistics. In this way we incorporate the MC statistical uncertainty within the final statistical uncertainty of the result.

- $N_{expected}(B \to D^0\pi, D^0 \to X/B \to D^0\pi) = 100$
  Sigma in the constraint = 3% (PDG) + 5.9% (coming from 290 reconstructed events in the final state) = 6.5%.

- $N_{expected}(B \to D^0 K, D^0 \to X) = 5$

  Sigma in the constraint = 9% (PDG) + 22% (coming from 20 reconstructed events in the final state) = 24%.

- $N_{expected}(B \to K^- \pi^+ \pi^-) = 11$

  Sigma in the constraint = 6.6% (PDG) + 5.5% (coming from 333 reconstructed events in the final state) = 8.6%.

- $N_{expected}(B^0 \to D_0^{*-} \ell^+ \nu_\ell) = 5$

  Sigma in the constraint = 40% (PDG) + 14% (coming from 54 reconstructed events in the final state) = 42%.

4.3 Maximum Likelihood fit

To separate the signal and background contributions and give an estimate of their yields we are going to perform a maximum likelihood fit.

A maximum likelihood fit [73] is a technique to estimate the value of some unknown parameters $\vec{\theta} = (\theta_1, \ldots, \theta_m)$, given a finite sample of data. Suppose to measure n times the random variable X[2], with *pdf* $f(X, \vec{\theta})$. The probability that a measurement will give the measured values $(x_1, x_2, \ldots, x_n)$ is given by the joint probability, the product of all *pdf*s. If the *pdf* and the parameter values are correct, we expect to have higher value of that probability for the measured data, while lower value in case of incorrect parameter values.

The likelihood function is defined as:

$$\mathcal{L} = \prod_{i=1}^{n} f(x_i | \vec{\theta}). \tag{4.1}$$

It has the same analytical expression of the joint probability with the difference that x_i are not the random variables X, but the values assumed by X in a specific measurement. The likelihood function is therefore only a function of the unknown parameters θ, while x_i are treated as fixed.

---

[2] If the outcome of an observation is characterized by several quantities, X can be a multi-dimensional variable.

In this way we can define the maximum likelihood (ML) estimators those which maximize the parameters $\theta_1, \ldots, \theta_m$, that are solutions of the equations:

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = 0, \qquad i = 1, \ldots, m.$$

This formalism is valid when the maximum is not at the boundary of the parameter range. When more than one local minimum exists, the highest one is taken. A convenient computational choice is rather to minimize the function $-2 \log \mathcal{L}$. In this analysis we use as minimization tool the `MINUIT` package [74, 75] [3].

There are several advantages to use the ML method: it is suitable for unbinned distributions and moreover the estimators are asymptotically normal, efficient [4], invariant [5] and sufficient[6].

In the cases when the *pdf* is given by the sum of various components (take for example the case of a sample given by the sum of a signal and a background components) and θ_j are the relative contribution of each component, $f(x_i|\theta)$ can be written as:

$$f(x_i|\theta) = \sum_{j=1}^{m} \theta_j f_j(x_i), \qquad (4.2)$$

where j runs over the number of contributions in the data. Every *pdf* $f_j(x_i)$ has to be known, and θ_j are constrained to have the total sum equal to one. We can write one of the parameter as a function of the other, as $\theta_m = 1 - \sum_{i=1}^{m-1} \theta_i$, reducing the degree of freedom to $m - 1$.

Moreover the likelihood function $\mathcal{L}$ is written as the product of the single-event likelihood $\mathcal{L}_i$, over the total number of events N of the sample:

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^{N} \mathcal{L}_i(\vec{\theta}). \qquad (4.3)$$

---

[3] The `MINUIT` package is composed by algorithms that vary the parameters θ until the convergence of the likelihood function and of the right covariance matrix is reached.

[4] An efficient estimator has minimum variance.

[5] If $\hat{\theta}$ is an invariant estimator of θ, then $g(\hat{\theta})$ is an estimator of $g(\theta)$.

[6] A sufficient estimator is defined in a unique way without other information.

4.3.1 The extended likelihood fit

There are cases where the total number n of observation is treated as a random variable, distributed with a Poisson function, with mean value ν. It generally happens when the outcomes of an experiment are the number n itself and the measured values $x_1, \ldots, x_n$. The likelihood function will be the product of the Poisson probability to find n and the likelihood function of eq. (4.1).

$$\mathcal{L}(\nu, \vec{\theta}) = \frac{\nu^n}{n!} e^{-\nu} \prod_{i=1}^{n} f(x_i|\vec{\theta}) = \frac{e^{-\nu}}{n!} \prod_{i=1}^{n} \nu f(x_i|\vec{\theta}) \qquad \boxed{4.4}$$

This is the general expression of the *extended likelihood function*.

Consider the case of this analysis, where we want to measure the relative fractions and the number of events of the decays in our data-sample . Note that the results of an extended and the corresponding not extended fit are the same, the only change is in the error values. In fact to use an extended fit allows to take into account the correlations among the total number and the fractions and to propagate the errors on the yields in the correct way.

From eq (4.4), taking the logarithm and dropping for terms not depending on the parameters, we obtain:

$$\log \mathcal{L}(\nu, \vec{\theta}) = -\nu + \sum_{i=1}^{n} \log \left(\sum_{j=1}^{m} \nu \theta_j f_j(x_i) \right),$$

and defining $\mu_i = \theta_i \nu$ as the expected number of events of type i, the log-likelihood function can be written as a function of the m parameters $\mu = (\mu_1, \ldots, \mu_m)$:

$$\log \mathcal{L}(\mu) = -\sum_{j=1}^{m} \mu_i + \sum_{i=1}^{n} \log \left(\sum_{j=1}^{m} \mu_j f_j(x_i) \right). \qquad \boxed{4.5}$$

Our sample is divided into four sub-sample,s corresponding to the favored, suppressed, positive and negative charges, each of them having their total number of events $\vec{n} = \{n_1, n_2, n_3, n_4\}$, with sum n_{tot}. The extended likelihood function will therefore be the product of a Poisson and a multinomial distribution. The first gives the probability of having n_{tot} total events, the second gives the probability of distributing the n_{tot} observations into four sub-sample. Of course the same description is valid in a general case of N sub-sample, as well as in the case of binned histograms.

The joint probability is therefore:

$$f(\vec{n}, \vec{\nu}) = \frac{\nu_{tot}^{n_{tot}} e^{-\nu_{tot}}}{n_{tot}!} \frac{n_{tot}!}{n_1! \dots n_4!} \left(\frac{\nu_1}{\nu_{tot}} \right)^{n_1} \dots \left(\frac{\nu_4}{\nu_{tot}} \right)^{n_4}$$

where $\vec{\nu}$ are the expected mean value of the $\vec{n}$, $\nu_{tot} = \sum_{i=1}^{4} \nu_i$ and $n_{tot} = \sum_{i=1}^{4} n_i$. This expression can be simplified in:

$$f(\vec{n}, \vec{\nu}) = \prod_{i=1}^{4} \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i}, \qquad (4.6)$$

which is equivalent to treat the number of events of each sub-sample as an independent Poisson random variable n_i with mean value ν_i.

The log-likelihood function becomes:

$$\log \mathcal{L}(\nu_{tot}|\vec{\theta}) = -\nu_{tot} + \sum_{i=1}^{4} n_i \log \nu_i(\nu_{tot}|\vec{\theta}). \qquad (4.7)$$

We are going to use this expression in the fit.

4.4 Likelihood anatomy

We perform the fit simultaneously on favored (*fav*) and suppressed (*sup*) samples. The expression of the extended Likelihood function is:

$$\mathcal{L} = \mathcal{P}_{fav+} \cdot \mathcal{L}_{fav+} \cdot \mathcal{P}_{fav-} \cdot \mathcal{L}_{fav-} \cdot \mathcal{P}_{sup+} \cdot \mathcal{L}_{sup+} \cdot \mathcal{P}_{sup-} \cdot \mathcal{L}_{sup-}, \qquad (4.8)$$

where $\mathcal{P}_i$ is the Poisson distribution of the single component (as discussed in Sect. 4.3.1), defined as:

$$\mathcal{P}_i = \frac{\mu^{N_i}}{N_i!} e^{-\mu}. \qquad (4.9)$$

The individual components of the likelihood are defined as in the following. We are using the convention to divide the components into "signal" and "background", with $(1 - b_i)$ and b_i the respective fractions. For convenience we write the *pdf* of the $B^- \to D^* \pi^-$ in the "signal" part because we verified from MC that this fraction remains constant in both samples.

The *pdf*s are written as functions of the mass (M) and of the particle identification variable κ. More detail will be given in the next sections.

$$\mathcal{L}_{fav+} = \prod_{i}^{N_{TOT}^{fav+}} \Big[(1 - b_{fav+}) \cdot \big(f_{\pi}^{fav+} \cdot pdf_{\pi}(M, \kappa) + \mathbf{c}^{+} \cdot f_{\pi}^{fav+} \cdot pdf_{D*}(M, \kappa)+$$

$$+ \big(1 - f_{\pi}^{fav+} - \mathbf{c}^{+} \cdot f_{\pi}^{fav+}\big) \cdot pdf_{K}(M, \kappa)\big) + b_{fav+} \cdot pdf_{comb}(M, \kappa)\Big]$$

$$\mathcal{L}_{fav-} = \prod_{i}^{N_{TOT}^{fav-}} \Big[(1 - b_{fav-}) \cdot \big(f_{\pi}^{fav-} \cdot pdf_{\pi}(M, \kappa) + \mathbf{c}^{-} \cdot f_{\pi}^{fav-} \cdot pdf_{D*}(M, \kappa)+$$

$$+ \big(1 - f_{\pi}^{fav-} - \mathbf{c}^{-} \cdot f_{\pi}^{fav-}\big) \cdot pdf_{K}(M, \kappa)\big) + b_{fav-} \cdot pdf_{comb}(M, \kappa)\Big]$$

$$\mathcal{L}_{sup+} = \prod_{i}^{N_{TOT}^{sup+}} \Big[(1 - b_{sup+}) \cdot \big(f_{\pi}^{sup+} \cdot pdf_{\pi}(M, \kappa) + \mathbf{c}^{+} \cdot f_{\pi}^{sup+} \cdot pdf_{D*}(M, \kappa)+$$

$$+ \big(1 - f_{\pi}^{sup+} - \mathbf{c}^{+} \cdot f_{\pi}^{sup+}\big) \cdot pdf_{K}(M, \kappa)\big) +$$

$$+ b_{sup+} \cdot \Big(f_{[X]\pi}^{+} \cdot pdf_{[X]\pi}(M, \kappa) + f_{[X]K}^{+} \cdot pdf_{[X]K} + f_{K\pi\pi}^{+} \cdot pdf_{K\pi\pi}(M, \kappa)+$$

$$f_{B^0}^{+} \cdot pdf_{B^0}(M, \kappa) + (1 - f_{[X]\pi}^{+} - f_{[X]K}^{+} - f_{K\pi\pi}^{+} - f_{B^0}^{+}) \cdot pdf_{comb}(M, \kappa)\Big)\Big]$$

$$\mathcal{L}_{sup-} = \prod_{i}^{N_{TOT}^{sup-}} \Big[(1 - b_{sup-}) \cdot \big(f_{\pi}^{sup-} \cdot pdf_{\pi}(M, \kappa) + \mathbf{c}^{-} \cdot f_{\pi}^{sup-} \cdot pdf_{D*}(M, \kappa)+$$

$$+ \big(1 - f_{\pi}^{sup-} - \mathbf{c}^{-} \cdot f_{\pi}^{sup-}\big) \cdot pdf_{K}(M, \kappa)\big) +$$

$$+ b_{sup-} \cdot \Big(f_{[X]\pi}^{-} \cdot pdf_{[X]\pi}(M, \kappa) + f_{[X]K}^{-} \cdot pdf_{[X]K} + f_{K\pi\pi}^{-} \cdot pdf_{K\pi\pi}(M, \kappa)+$$

$$f_{B^0}^{-} \cdot pdf_{B^0}(M, \kappa) + (1 - f_{[X]\pi}^{-} - f_{[X]K}^{-} - f_{K\pi\pi}^{-} - f_{B^0}^{-}) \cdot pdf_{comb}(M, \kappa)\Big)\Big]$$

The parameters b_{fav+}, b_{fav-}, b_{sup+} and b_{sup-} are the fractions of the background for each mode and charge. In the favored likelihood we consider only the combinatorial background, for which we use a single pdf_{comb} for both positive and negative charges. For the suppressed we are considering the combinatorial (with the same pdf of the favored sample) and the physical backgrounds $B \to K\pi\pi$ in three-body decay (of which the fraction is $f_{K\pi\pi}^{\pm}$ and $pdf_{K\pi\pi}$), $B^- \to D(\to X)\pi^-$ (of which the fraction is $f_{[X]\pi}^{\pm}$ and $pdf_{[X]\pi}$), $B^- \to D(\to X)K^-$ (of which the fraction is $f_{[X]K}^{\pm}$ and $pdf_{[X]K}$), and $B^0 \to D_0^{*-}l^+\nu_l$ (of which the fraction is $f_{B^0}^{\pm}$ and pdf_{B^0}).

For the signals, $f_\pi^{fav,sup,\pm}$ is the fraction of $B^- \to D\pi^-$ for favored, suppressed, positive and negative charges, $\mathbf{c}^\pm$ is one of the common parameter of favored and suppressed likelihoods and corresponds to the ratio of the $B^- \to D^*\pi^-$ over $B^- \to D\pi^-$. The simultaneous fit allows us to take advantage of the favored channel with more statistics to constrain the common parameter $\mathbf{c}$ in a consistent way.

The fraction of $B^- \to DK^-$ is written as $(1 - f_\pi - \mathbf{c} \cdot f_\pi)$, so it is not explicitly fitted, but it is calculated from the fraction of $B^- \to D\pi^-$ and $B^- \to D^{*0}\pi^-$. This is fine for the fit itself, but makes the calculation of the suppressed $B^- \to DK^-$ significance, where that fraction has to be fixed to zero, hard to converge to realistic values. For this reason we decided to apply this change of variables in the suppressed part of the likelihood:

$$1 - f_\pi^{sup} - \mathbf{c} \cdot f_\pi^{sup} = f_K^{sup}$$

$$f_\pi^{sup} = 1 - f_K^{sup} - \mathbf{c} \cdot f_\pi^{sup},$$

where the latter expression can be simplified in: $f_\pi^{sup} = (1 - f_K^{sup})/(1 + \mathbf{c})$.

In this way the formal expressions of the suppressed likelihood are:

$$
\begin{aligned}
\mathcal{L}_{sup+} = \prod_i^{N_{TOT}^{sup+}} \Big[& (1 - b_{sup+}) \cdot \big((1 - f_K^{sup+})/(1 + \mathbf{c}^+) \cdot pdf_\pi(M, \kappa) + \\
& + \mathbf{c}^+ f_\pi^{sup+} \cdot pdf_{D*}(M, \kappa) + f_K^{sup+} \cdot pdf_K(M, \kappa) \big) + \\
& + b_{sup+} \cdot \big(f_{[X]\pi}^+ \cdot pdf_{[X]\pi}(M, \kappa) + f_{[X]K}^+ \cdot pdf_{[X]K} + f_{K\pi\pi}^+ \cdot pdf_{K\pi\pi}(M, \kappa) + \\
& f_{B^0}^+ \cdot pdf_{B^0}(M, \kappa) + (1 - f_{[X]\pi}^+ - f_{[X]K}^+ - f_{K\pi\pi}^+ - f_{B^0}^+) \cdot pdf_{comb}(M, \kappa) \big) \Big]
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{L}_{sup-} = \prod_i^{N_{TOT}^{sup-}} \Big[& (1 - b_{sup-}) \cdot \big((1 - f_K^{sup-})/(1 + \mathbf{c}^-) \cdot pdf_\pi(M, \kappa) + \\
& + \mathbf{c}^- f_\pi^{sup-} \cdot pdf_{D*}(M, \kappa) + f_K^{sup-} \cdot pdf_K(M, \kappa) \big) + \\
& + b_{sup-} \cdot \big(f_{[X]\pi}^- \cdot pdf_{[X]\pi}(M, \kappa) + f_{[X]K}^- \cdot pdf_{[X]K} + f_{K\pi\pi}^- \cdot pdf_{K\pi\pi}(M, \kappa) + \\
& f_{B^0}^- \cdot pdf_{B^0}(M, \kappa) + (1 - f_{[X]\pi}^- - f_{[X]K}^- - f_{K\pi\pi}^- - f_{B^0}^-) \cdot pdf_{comb}(M, \kappa) \big) \Big]
\end{aligned}
$$

The likelihood function is the same as before, so it will give the same result as before, and, as a cross-check, we verified that all yields and observable results are

the same. The advantage is now that the fraction of $B^- \to DK^-$ can be directly fixed to zero allowing an easier calculation of the significance. This has been chosen as the final fit configuration.

As already explained, the *pdf*s are functions of the mass M, reconstructed in the $(K\pi)\pi$ hypothesis and of Particle Identification (PID) through the kaoness variable κ applied to the track from B. Each *pdf* characterizes a different decay, but they are common for favored and suppressed likelihoods.

In the next sections we will consider piece by piece all *pdf*s, and describe the parameterizations used for mass and PID variables.

4.5 Mass templates

4.5.1 Signal mass template

Monte Carlo samples of $B^- \to D\pi^-$ and $B^- \to DK^-$ have been generated including the effect of radiations in the final state (see Sect. 2.10).

The mass line shape is parameterized using the following asymmetric pdf:

$$
\begin{aligned}
pdf(m; \vec{\theta}) = f_{fsr} \cdot & \left(\frac{1}{norm} \cdot e^{b_{fsr} \cdot (m - (mu + \Delta))} \cdot erfc(c_{fsr} \cdot (m - (mu + \Delta))) \right) \\
+ (1 - f_{fsr}) \cdot & \left(f_1 \frac{1}{\sigma_1 \cdot s \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m - (mu + \Delta)}{\sigma_1 \cdot s} \right)^2} + f_2 \frac{1}{\sigma_2 \cdot s \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m - (mu + \Delta)}{\sigma_2 \cdot s} \right)^2} \right. \\
+ f_3 & \left. \frac{1}{\sigma_3 \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m - (mu + \Delta)}{\sigma_3} \right)^2} \right)
\end{aligned}
\tag{4.10}
$$

where f_{fsr} is the contribution of the radiative tail fraction. f_1, f_2, f_3 and $\sigma_1, \sigma_2, \sigma_3$ are respectively the fractions and the widths of the three gaussians ($f_1 + f_2 + f_3 = 1$). Δ is a mass scale parameter that is left free to be determined by the fit. We also introduced a free scale parameter s in the likelihood fit, that multiplies the width of the three gaussians, to account for differences between data and MC. Fig. 4.11 show the invariant mass distribution of the simulated $B^- \to D\pi^-$ events, while Fig. 4.12 show the invariant mass distribution of the simulated $B^- \to DK^-$ events, both with the superimposition of the fit function (4.10).

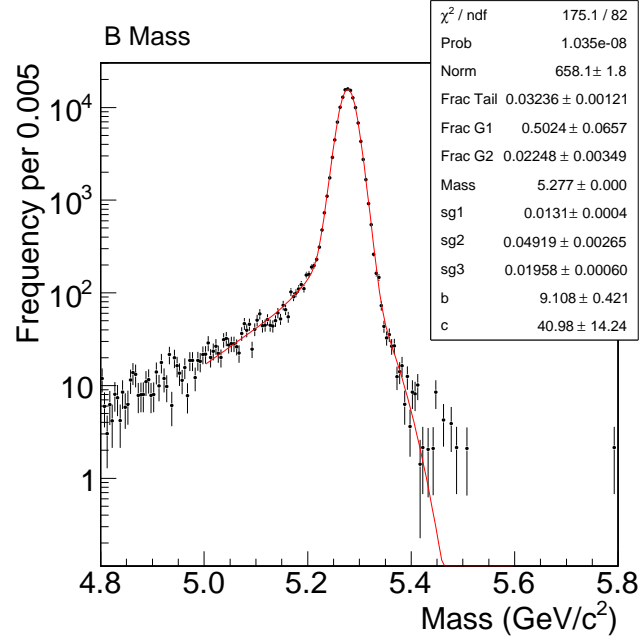| χ^2 / ndf	175.1 / 82
Prob	1.035e-08
Norm	658.1± 1.8
Frac Tail	0.03236 ± 0.00121
Frac G1	0.5024 ± 0.0657
Frac G2	0.02248 ± 0.00349
Mass	5.277 ± 0.000
sg1	0.0131± 0.0004
sg2	0.04919 ± 0.00265
sg3	0.01958 ± 0.00060
b	9.108 ± 0.421
c	40.98 ± 14.24

Figure 4.11: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to D\pi^-$ events on logarithmic scale. We overlaided the fit function (4.10).
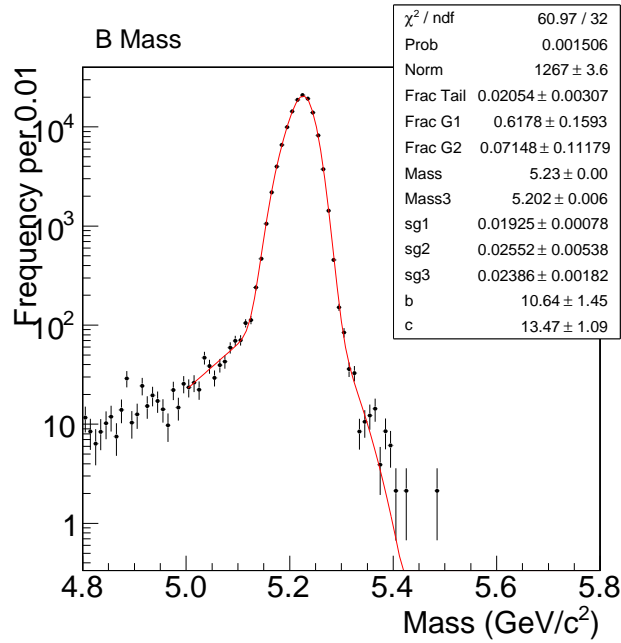

χ^2 / ndf	60.97 / 32
Prob	0.001506
Norm	1267 ± 3.6
Frac Tail	0.02054 ± 0.00307
Frac G1	0.6178 ± 0.1593
Frac G2	0.07148 ± 0.11179
Mass	5.23 ± 0.00
Mass3	5.202 ± 0.006
sg1	0.01925 ± 0.00078
sg2	0.02552 ± 0.00538
sg3	0.02386 ± 0.00182
b	10.64 ± 1.45
c	13.47 ± 1.09

Figure 4.12: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to DK^-$ events on logarithmic scale. We overlaided the fit function (4.10).

4.5.2 Physics background

$B^- \to \overline{D}^{*0} \pi^-$

As defined in section 4.1.1, the fit window is [5.17, 6.5] GeV/c^2 to have only the $B^- \to \overline{D}^{*0} \pi^-$ contribution as physics background in the favored sample. In Fig. 4.13 the invariant mass distribution of simulated $B^- \to D^{0*} \pi^-$ (log-scale) events, is shown. The distribution has been fitted in the window [5.14,5.6] GeV/c^2 with the sum of three gaussians plus an exponential.
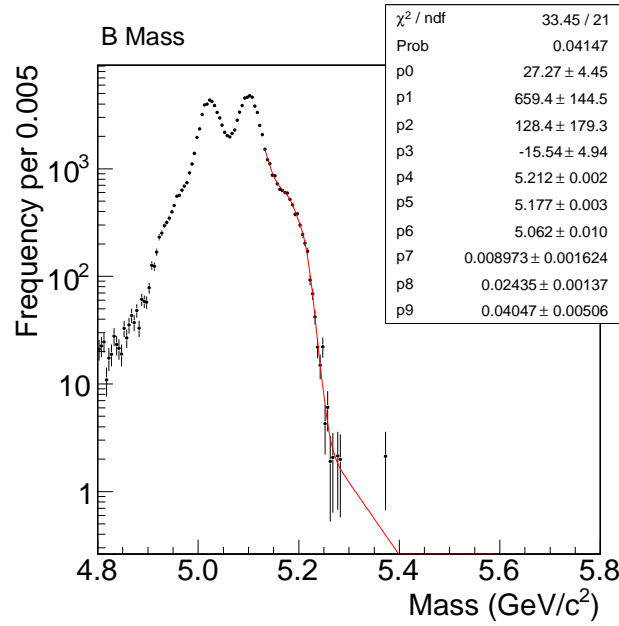

χ^2 / ndf	33.45 / 21
Prob	0.04147
p0	27.27 ± 4.45
p1	659.4 ± 144.5
p2	128.4 ± 179.3
p3	-15.54 ± 4.94
p4	5.212 ± 0.002
p5	5.177 ± 0.003
p6	5.062 ± 0.010
p7	0.008973 ± 0.001624
p8	0.02435 ± 0.00137
p9	0.04047 ± 0.00506

Figure 4.13: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to \overline{D}^{0*} \pi^-$ events on logarithmic scale. We overlaided the fit function (sum of three gaussians and an exponential).

Three-body decay

Fig. 4.14 shows the invariant mass distribution of $B^- \to K^- \pi^+ \pi^-$ from MC simulation. The decay is reconstructed as suppressed $B^- \to D\pi^-$. The distribution has been fitted with the sum of two gaussians.

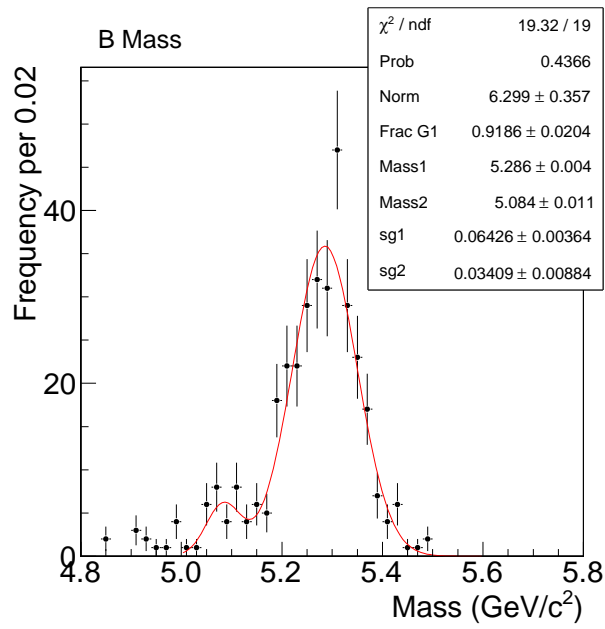| χ^2 / ndf	19.32 / 19
Prob	0.4366
Norm	6.299 ± 0.357
Frac G1	0.9186 ± 0.0204
Mass1	5.286 ± 0.004
Mass2	5.084 ± 0.011
sg1	0.06426 ± 0.00364
sg2	0.03409 ± 0.00884

Figure 4.14: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to K^- \pi^+ \pi^-$ events. We overlaided the fit function (sum of two gaussians).

$B^- \to D(\to X)\pi^-$

The $B^- \to D(\to X)\pi^-$ decay, where X are modes other than $K^+\pi^-$ is reconstructed as suppressed $B^- \to D\pi^-$. Fig. 4.15 shows the invariant mass distribution of this background with superimposed the fit function (a Pearson function of IV type) [7].
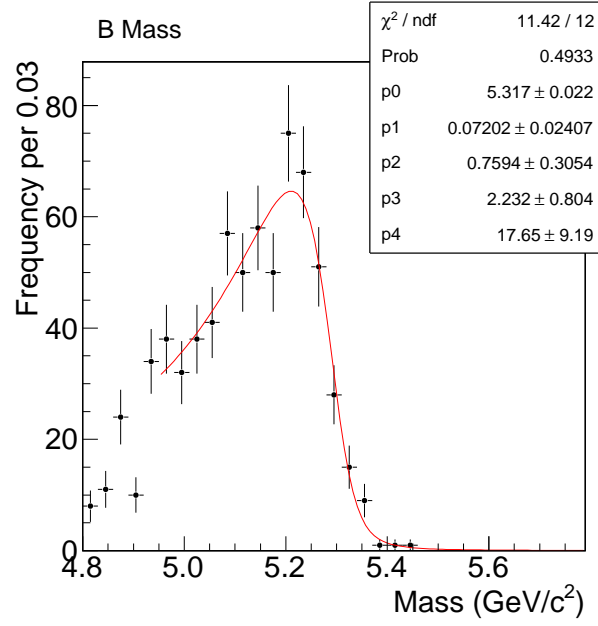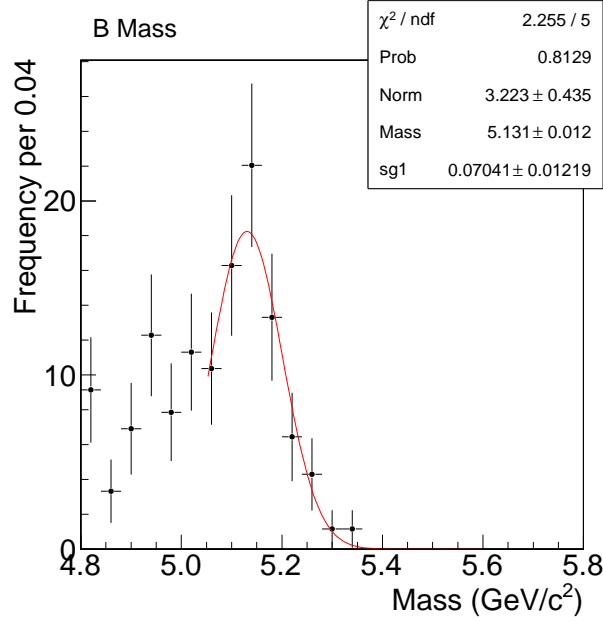

Figure 4.15: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to D(\to X)\pi^-$ events. We overlaided the fit function (a Pearson function of IV type).

$B^- \to D(\to X)K^-$

The $B^- \to D^0 K^-$ decay with $D^0 \to X$ is reconstructed as $B^- \to D_{sup}\pi^-$. Fig. 4.16 shows the invariant mass distribution of this background with superimposed the fit function (one gaussian).

---

[7]The Pearson function of the IV type is described by

$$f(x)dx = k\left[1 + \left(\frac{x-\lambda}{a}\right)^2\right]^{-m} \exp\left[-\nu \tan^{-1}\left(\frac{x-\lambda}{a}\right)\right]dx \qquad (m > 1/2)$$

χ^2 / ndf	2.255 / 5
Prob	0.8129
Norm	3.223 ± 0.435
Mass	5.131 ± 0.012
sg1	0.07041 ± 0.01219

Figure 4.16: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^- \to D(\to X)K^-$ events. We overlaided the fit function (one gaussian).

$B^0 \to D_0^{*-} \ell^+ \nu_\ell$

The $B^0 \to D_0^{*-} \ell^+ \nu_\ell$, with $D_0^{*-} \to \bar{D}^0 \pi^-$, with $\bar{D}^0 \to K^+ \pi^-$ is reconstructed as $B^- \to D_{sup}\pi^-$. Fig. 4.17 shows the invariant mass distribution of this background with superimposed the fit function (one gaussian).

4.5.3 Combinatorial background

We describe the combinatorial background mass shape with an exponential. The slope and the normalization of the exponential are left free in the main fit. We use the same shape for both samples.
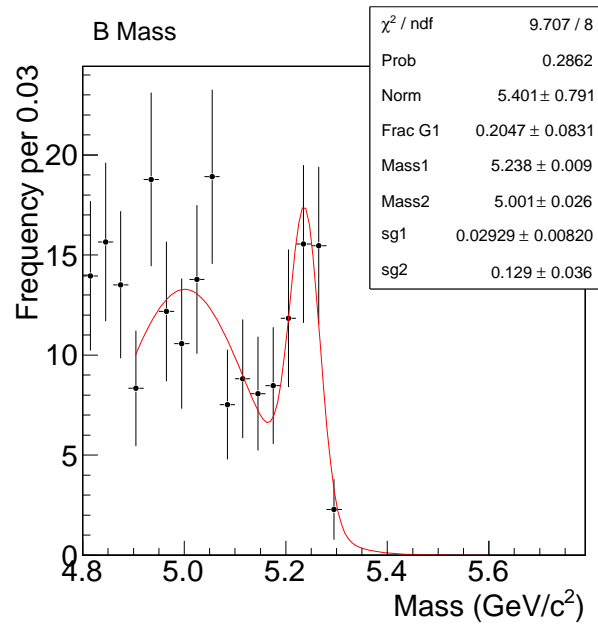
Figure 4.17: MC simulation, PHOTOS turned on: invariant mass distribution of the simulated $B^0 \to D_0^{*-}\ell^+\nu_\ell$, with $D_0^{*-} \to \bar{D}^0\pi^-$, with $\bar{D}^0 \to K^+\pi^-$ events. We overlaided the fit function (one gaussian).

4.6 PID parameterization (signal and background)

To distinguish between $B^- \to DK^-$ and $B^- \to D\pi^-$ modes, particle identification is applied to the track from the B. Since our selection requires track with $p_T > 2$ GeV/c^2, we can use the standard calibration of the dE/dx and Universal Curve (UC) (see Sect. 2.3.2).

The PID information can be represented by the single observable κ, defined in Sect. 3.3:

$$\kappa = \frac{\frac{dE}{dx}_{meas} - \frac{dE}{dx}_{exp}(\pi)}{\frac{dE}{dx}_{exp}(K) - \frac{dE}{dx}_{exp}(\pi)}. \tag{4.11}$$

where $\frac{dE}{dx}_{meas}$ is the measured energy lost by the charged particle in the drift chamber volume and $\frac{dE}{dx}_{exp}(\pi, K)$ is the corresponding expected value for a given particle hypothesis. The distribution of κ is momentum-independent with good approximation for tracks with momentum greater than 2 GeV/c^2, and this provides the important advantage of not having to include momentum information in the likelihood [8]. Moreover it is independent from the mass, allowing to separate the contribution of the mass and PID by simply multiplying the *pdf*s without considering the correlations.

We determined the *pdf*s of the κ variable directly from our data-sample. We selected kaons and pions from the D of the favored B^- decay, requiring both the pion and the kaon from D to be trigger tracks. Parameterizing this variable in our sample allows to take into account any systematic effect due to the fact that the UC are not extracted in our sample and to avoid sample dependence effects. This is because the PID templates are constructed in the same data that we fit.

To take into account of the different kinematic of D and B decays, we reweighted the transverse momentum distribution (p_T) of kaons and pions from D with the p_T of kaons and pions from B^-. The p_T distributions of kaons and pions from D are taken from data, while the p_T distributions of kaons and pions from B^- are taken from MC (in Sect. 4.1 Fig. 4.6 shows the good agreement between data and MC for

---

[8]In fact, whenever the templates used in a multi-component fit depend on additional observables, one should always use complete Likelihood expression, including the explicit distributions of all observables for all classes of events. In our case this would mean to consider the momentum in all the likelihood terms. Otherwise the likelihood is not correct and can generate bias in the result [76]

the p_T distribution of the track from B). This reweighting is necessary due to the fact that the kaon-pion separation is momentum dependent so we can parameterize the κ distribution using kaons and pions from D only after the p_T reweighting.

Fig. 4.18 shows the κ distribution, for pions and for kaons after the reweighting. The two distribution have been fitted using the sum of three gaussians. We are going to use this parameterization in the maximum likelihood fit.
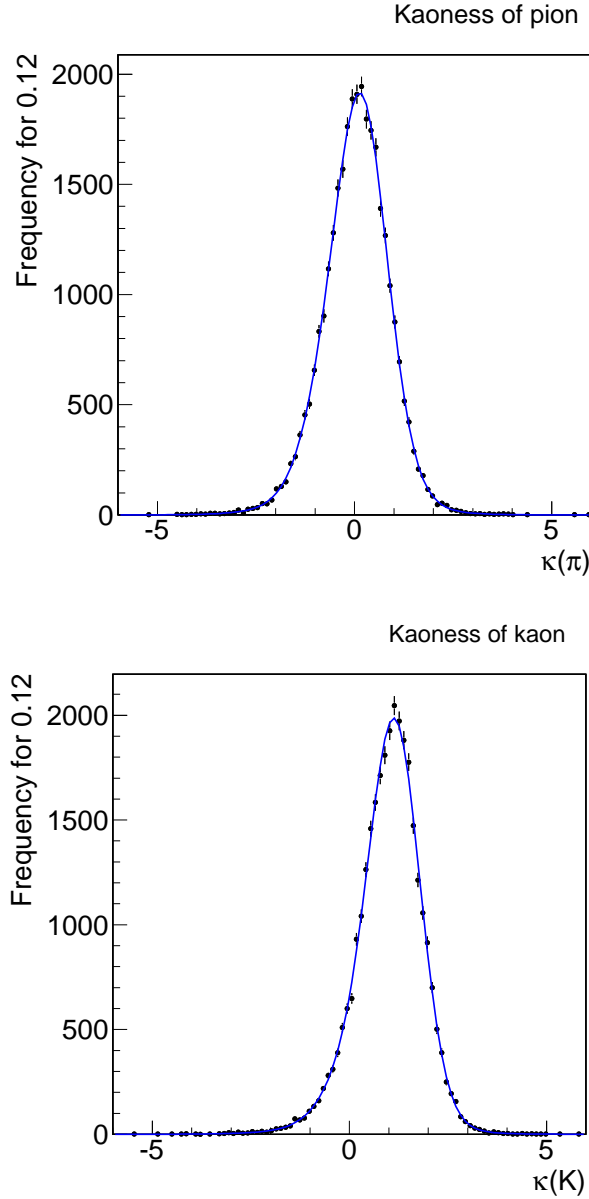
Kaoness of pion

Kaoness of kaon

Figure 4.18: Top: Kaoness distribution of pions from D of the favored B^- decay. Bottom: Kaoness distribution of kaons from D of the favored B^- decay. The fit functions are superimposed.

For the combinatorial background we assume that it is composed only by pions and kaons: the corresponding PID *pdf* will have the form:

$$a \cdot pdf_\pi(\kappa) + (1 - a) \cdot pdf_K(\kappa) \qquad (4.12)$$

where a is the fraction of pions in the combinatorial background. This fraction is left free to float in the maximum likelihood fit.

In the next chapter we will apply the fit to our samples and we will see how we obtain the final results.

Extraction of the results

In this chapter we will describe the extraction of the fit results and the studies performed to check for the possible presence of biases.

In addition we will describe how we correct the results for detector acceptance.

5.1 Fit results

The fit is simultaneously applied on both favored and suppressed samples, as selected with the optimized criteria listed in Sect. 3.3.5 (Fig. 3.15). The fit ranges are $5.17 \leq M(D\pi) \leq 6.5$ GeV/c^2 for the invariant mass of the B candidate, and $-3 \leq \kappa \leq 3$ for the kaoness of the track from B. The motivation for the first range is given in Sect. 4.1.1, the latter is chosen to avoid contributions from events with not well reconstructed or identified tracks, that fall in the tails of the kaoness.

From the fit we obtain the total number of events for each of the four data sub-samples and the fractions of signal and background contributions. Using the symbols introduced in Sect. 4.4, the yields of the events (N_i^j, where i is the mode and j is the sample) can be written as:

$$
\begin{aligned}
N_\pi^{fav\pm} &= N_{TOT}^{fav\pm} \cdot (1 - b_{fav\pm}) \cdot f_\pi^{fav\pm} \\
N_K^{fav\pm} &= N_{TOT}^{fav\pm} \cdot (1 - b_{fav\pm}) \cdot (1 - f_\pi^{fav\pm} - \mathbf{c}^\pm \cdot f_\pi^{fav\pm}) \\
N_\pi^{sup\pm} &= N_{TOT}^{sup\pm} \cdot (1 - b_{sup\pm}) \cdot (1 - f_K^{fav\pm})/(1 + \mathbf{c}^\pm) \\
N_K^{sup\pm} &= N_{TOT}^{sup\pm} \cdot (1 - b_{sup\pm}) \cdot f_K^{fav\pm}
\end{aligned}
\tag{5.1}
$$

5. Extraction of the results

In a similar way we can write the yields for each class of events.

The error on the yields is obtained propagating the errors on the fractions and on the total number of fitted events, according to the formula:

$$Var(y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} \mid_{\vec{c}=\vec{\mu}} (Var(\vec{x}))_{i,j} \qquad \boxed{5.2}$$

In Table 5.1 we report the fractions and yields of the signal events, in Table 5.2 the fractions and yields of background events, and in Table 5.3 the fit results for the remaining parameters.

From Table 5.1 we can see that the suppressed signals are different from zero. To evaluate their significance with respect to the background-only hypothesis we repeat the fit on data fixing to zero the fractions of suppressed $B^- \to DK^-$ ($B^- \to D\pi^-$) and calculate the difference of $-2\log\mathcal{L}$ (DLL) between this fit and the central fit. We separately fix to zero the fractions $B^+ \to D(\to K^-\pi^+)K^+$ and $B^- \to D(\to K^+\pi^-)K^-$ to evaluate the $B^- \to DK^-$ significance and then the fractions $B^+ \to D(\to K^-\pi^+)\pi^+$ and $B^- \to D(\to K^+\pi^-)\pi^-$ to evaluate the $B^- \to D\pi^-$ significance. The distribution of $-2\log\mathcal{L}$ approximately follows a χ^2 distribution with m degrees of freedom given by the number of parameters fixed to zero. We can therefore convert the value of $-2\log\mathcal{L}$ into a p-value, the probability to obtain the same or a worse result if repeating the experiment, and then into a number of sigma [1].

We obtain a $-2\Delta\log\mathcal{L}$ value of 14.1 for the $B^- \to DK^-$, corresponding to a statistical significance of 3.3σ and a value of 16.3 for the $B^- \to D\pi^-$, corresponding to a statistical significance of 3.6σ. These numbers are encouraging that we may have an evidence for the presence of such decay modes in our sample - however, since this is based on statistical uncertainties only, a definite conclusion cannot be reached before examining the effects of systematic uncertainties (Sect. 6.2).

In Table 5.2 the yields of the background components are all in agreement with expected values from Sect. 4.2.

In order to see if these value returned by the fit well describe the data, we do some checks, described in the following sections.

---

[1] To convert a p-value (α) in number of sigma (n), we applied the formula according to PDG convention: $n = \sqrt{2} \cdot Erf^{-1}(1-\alpha)$.

Decay	fraction	yield
$B^+ \to \overline{D}^0 \pi^+$	(f_π^{fav+}) 0.904 ± 0.004	9882 ± 103
$B^- \to D^0 \pi^-$	(f_π^{fav-}) 0.893 ± 0.004	9892 ± 103
$B^+ \to \overline{D}^0 K^+$	(f_K^{fav+}) 0.063 ± 0.004	694 ± 39
$B^- \to D^0 K^-$	(f_K^{fav-}) 0.069 ± 0.004	767 ± 41
$B^+ \to \overline{D}^0 \pi^+$	(f_π^{sup+}) 0.44 ± 0.12	24 ± 9
$B^- \to D^0 \pi^-$	(f_π^{sup-}) 0.88 ± 0.20	31 ± 10
$B^+ \to \overline{D}^0 K^+$	(f_K^{sup+}) 0.55 ± 0.13	29 ± 9
$B^- \to D^0 K^-$	(f_K^{sup-}) 0.08 ± 0.20	3 ± 8

Table 5.1: Results of signal events from the maximum likelihood fit.

Decay	fraction	yield
$B^+ \to D^{0*} \pi^+$ fav	$(\mathbf{c}^+)$ 0.034 ± 0.003	355 ± 27
$B^- \to \overline{D}^{0*} \pi^-$ fav	$(\mathbf{c}^-)$ 0.041 ± 0.003	415 ± 29
$B^+ \to D^{0*} \pi^+$ sup	$(\mathbf{c}^+)$ 0.034 ± 0.003	1 ± 1
$B^- \to \overline{D}^{0*} \pi^-$ sup	$(\mathbf{c}^-)$ 0.041 ± 0.003	1 ± 1
$B^+ \to D(\to X)\pi^+$	$(f_{[X]\pi}^+)$ 0.16 ± 0.01	50 ± 7
$B^- \to D(\to X)\pi^-$	$(f_{[X]\pi}^-)$ 0.15 ± 0.01	50 ± 7
$B^+ \to D(\to X)K^+$	$(f_{[X]K}^+)$ 0.005 ± 0.001	1 ± 1
$B^- \to D(\to X)K^-$	$(f_{[X]K}^-)$ 0.005 ± 0.001	2 ± 2
$B^+ \to K^+ \pi^- \pi^+$	$(f_{K\pi\pi}^+)$ 0.017 ± 0.001	5 ± 2
$B^- \to K^- \pi^- \pi^+$	$(f_{K\pi\pi}^-)$ 0.016 ± 0.001	6 ± 2
$B^0 \to D_0^{*-} \ell^+ \nu_\ell$	$(f_{B^0}^+)$ 0.007 ± 0.003	2 ± 2
$B^0 \to D_0^{*+} \ell^- \bar{\nu}_\ell$	$(f_{B^0}^-)$ 0.007 ± 0.003	3 ± 2

Table 5.2: Results of background events from the maximum likelihood fit.

Parameter	Value
number (fav pos)	$(N_{TOT}^{fav+})\ 11217 \pm 106$
number (fav neg)	$(N_{TOT}^{fav-})\ 11362 \pm 107$
number (sup pos)	$(N_{TOT}^{sup+})\ 378 \pm 19$
number (sup neg)	$(N_{TOT}^{sup-})\ 371 \pm 19$
background fraction (fav pos)	$(b_{fav+})\ 0.026 \pm 0.002$
background fraction (fav neg)	$(b_{fav-})\ 0.025 \pm 0.002$
background fraction (sup pos)	$(b_{sup+})\ 0.86 \pm 0.03$
background fraction (sup neg)	$(b_{sup-})\ 0.90 \pm 0.03$
f_π in the comb. back (fav & sup)	0.69 ± 0.03
slope comb. back (fav & sup)	-2.6 ± 0.1
Δ	$(3 \pm 1) \cdot 10^{-4}$ GeV/c^2
Scale	(s) 1.107 ± 0.008

Table 5.3: Other parameters resulting from the fit.

5.2 Fit projections

An important immediate check that must be done to verify the reliability of the fit results is to compare the probability distributions predicted by the likelihood expression, after fixing the parameters to their fitted values, to the actual distributions in the data. In practice, this is more easily done by comparing one- or two-dimensional "projections" of the fit on individual observables. More precisely this is done by comparing a *"projection"* of the generic probability density function $pdf(x, \vec{y}|\vec{\theta})$ on the observable x, defined as the plot of the function $\int_{\vec{y}} pdf(x, \vec{y}|\vec{\theta}) d\vec{y}$ overlaid on the experimental data after an appropriate normalization. The projections allow to detect possible small discrepancies between the real data distributions and the model, or possible failures of the numerical algorithms in finding the correct maximum of the likelihood.

Figs. [5.1-5.7] show the fit projections on mass and kaoness variables. Points are data and solid lines are the fit projections. For each projection the difference between the total projection and data, divided by the statistical uncertainty in each bin, is shown. In particular we can see on Fig. 5.1 the projection on the mass sub-range [5.17, 5.6] GeV/c² of the favored sample, for positive (left) and negative (right) charges. They are shown in this sub-range to check the agreement of the fit for the signal contribution part. In Fig. 5.2 we can see the full range projection, in logarithmic scale, to visualize the agreement also for the combinatorial background in higher values of the mass.

Analogous distributions are shown for the suppressed sample on Figs. 5.3 and 5.4. The physical backgrounds are added together in a single component. We can clearly see by eye a strong asymmetry for the $B^- \to DK^-$ mode (red shape), which has almost null yield of negative charges, as can be also seen in Table 5.1. In Sect. 6.2 we will evaluate the significance of this asymmetry, to give a quantitative estimate of this quantity.
Fig. 5.5 shows in logarithmic scale the background contributions for the suppressed sample.

Finally on Figs. 5.6 and 5.7 we can see the projections on the kaoness variable

for favored and suppressed samples, respectively.

The agreement in both mass and kaoness variables, between data and fit functions, appears to be very satisfactory.

In the next Section we will discuss further tests of the intrinsic consistency of the fit, as well as the accuracy of the error estimates.
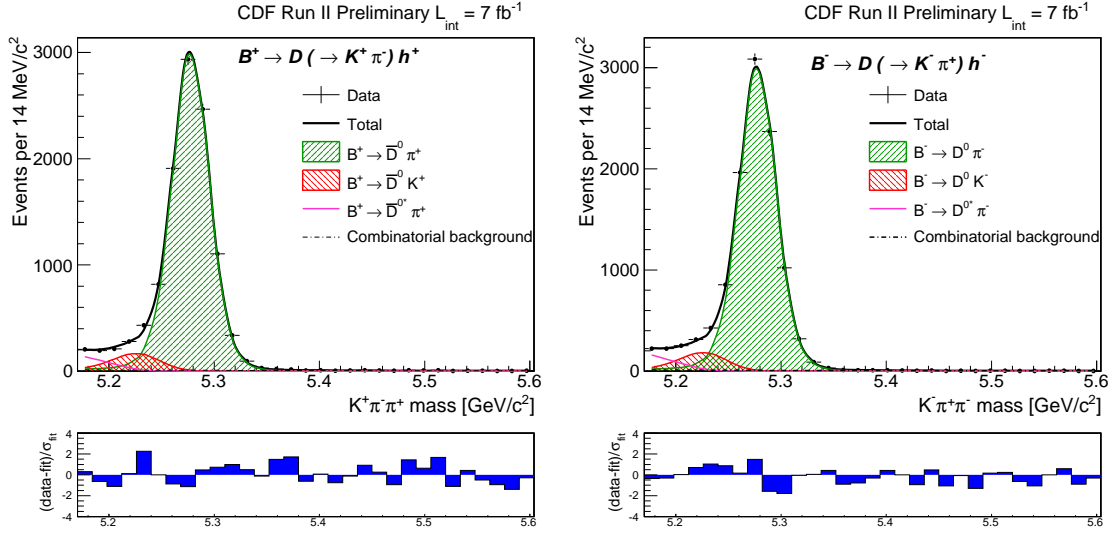

Figure 5.1: Fit projection in mass sub-range [5.17,5.6] GeV/c² for favored modes, positive (left) and negative charges (right).
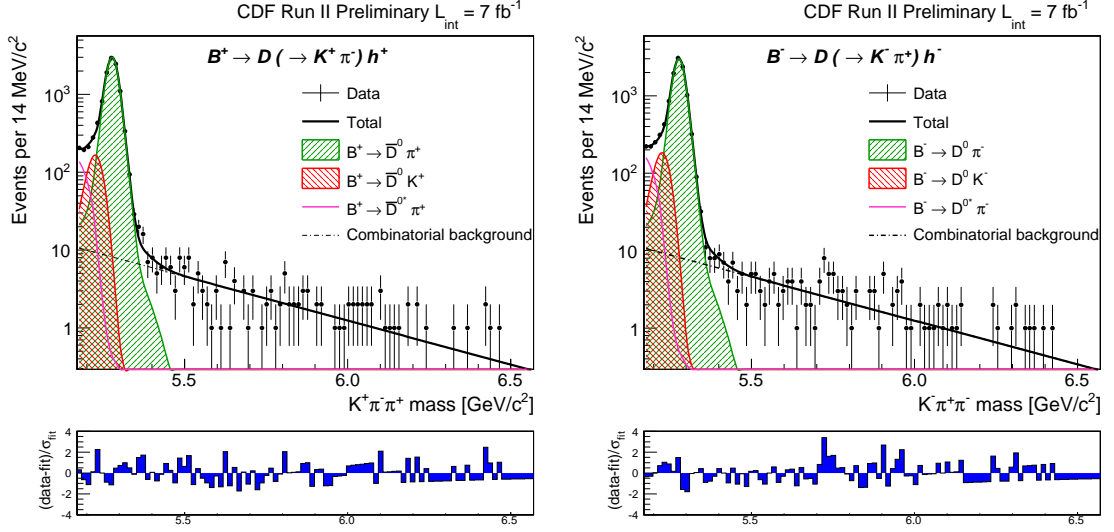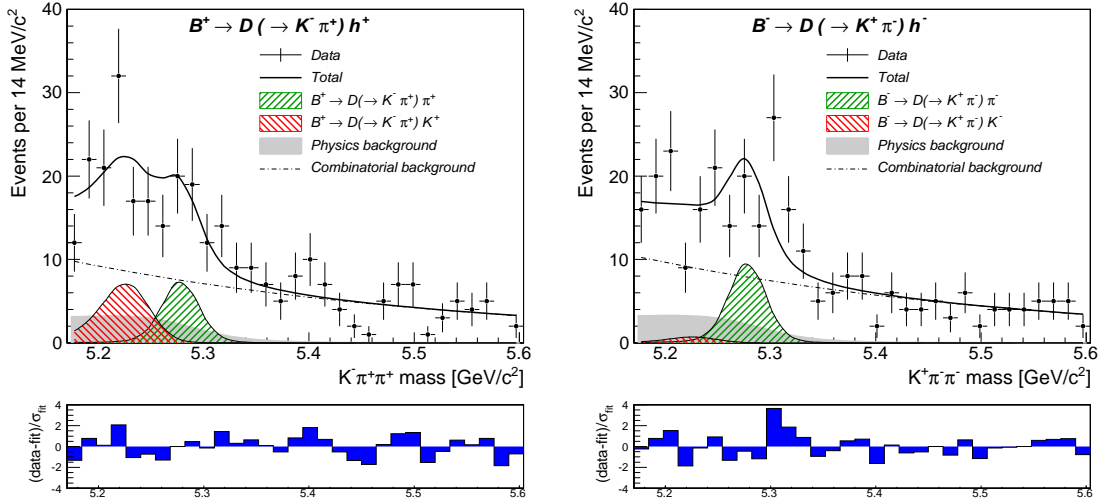
Figure 5.2: Fit projection in the full mass range [5.17,6.5] GeV/c² for favored modes, positive (left) and negative charges (right). The logarithmic scale allows to see the fit agreement also for the combinatorial background in the right tail of the plot.

Figure 5.3: Fit projection in mass sub-range [5.17,5.6] GeV/c² for suppressed mode, positive (left) and negative charges (right), with the physics background contributions summed together.
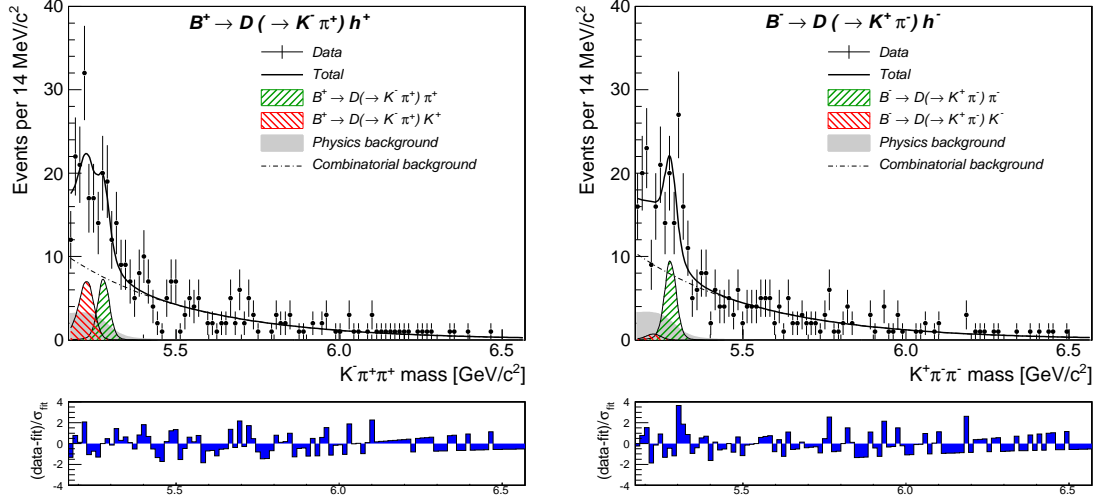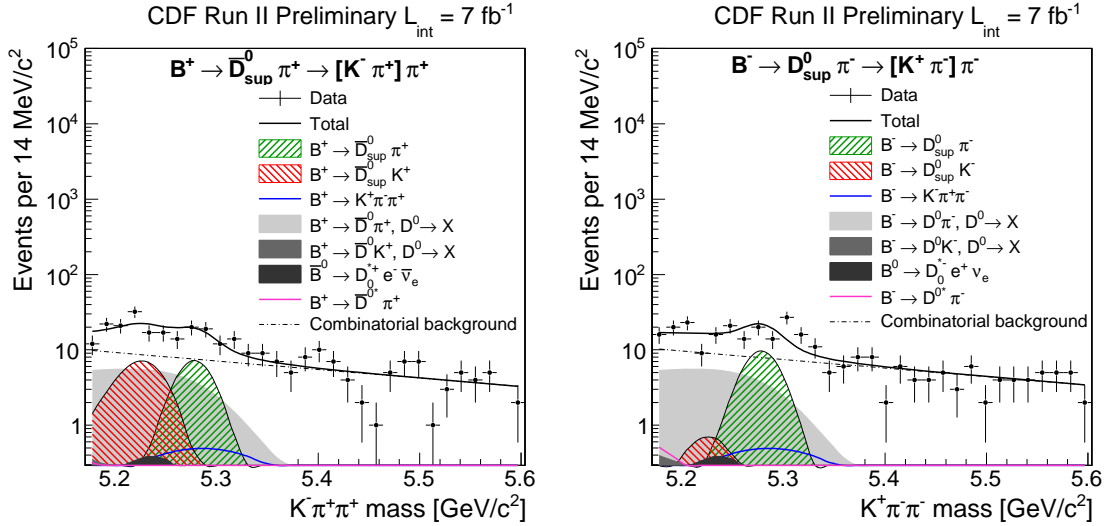
Figure 5.4: Fit projection in the full mass range [5.17,6.5] GeV/c^2 for suppressed mode, positive (left) and negative charges (right).

Figure 5.5: Fit projection in mass sub-range [5.17,5.6] GeV/c^2 for suppressed mode, positive (left) and negative charges (right), in logarithmic scale to see each backgrounds contribution.
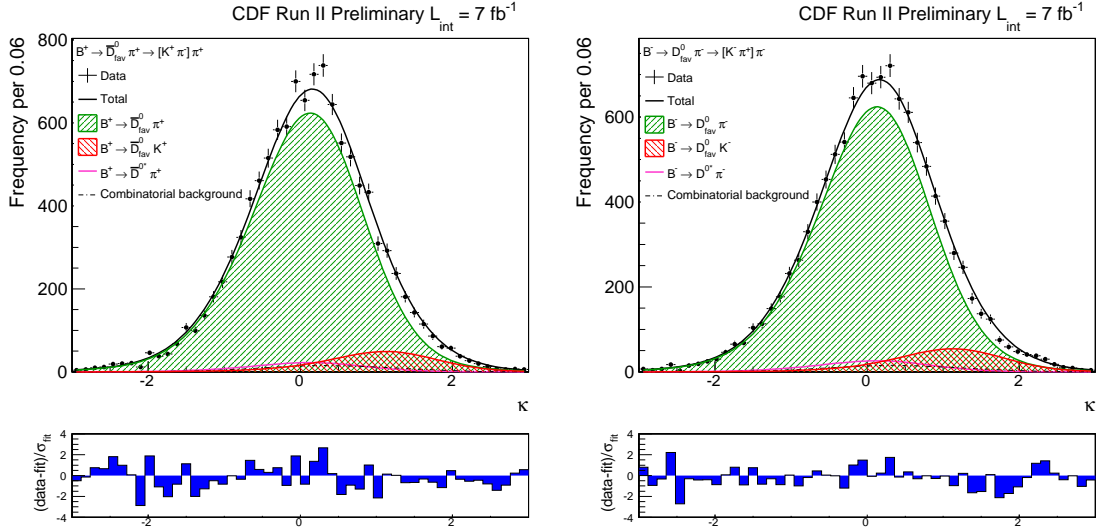
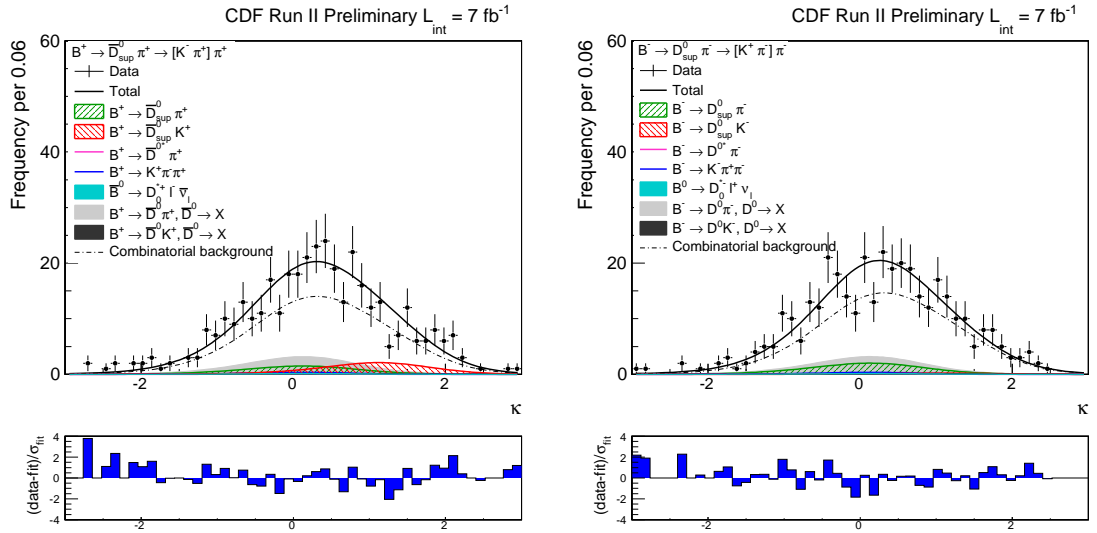Figure 5.6: Fit projection in the kaoness range [-3,3] for favored mode, positive (left) and negative charges (right).

Figure 5.7: Fit projection in the kaoness range [-3,3] for suppressed mode, positive (left) and negative charges(right).

5.3 Checks for fit consistency

Maximum likelihood fit based on samples of finite statistics can exhibit intrinsic biases. To investigate this possibility we apply the fit to many simulated pseudo-experiments and we evaluate the distribution of differences between fitted results from simulation and true values.

We generate 1000 pseudo-experiments, each of them consisting of simulated signal and background events with their proper *pdf*, product of the mass and PID *pdf*s. The number of total events fluctuates according to a Poisson distribution and the number of events of each favored, suppressed, positive and negative sub-samples fluctuate with a multinomial distribution. The mean values of these distributions are taken from data (according to the Table 5.1-5.3). This choice properly accounts for the statistical fluctuations of signal and background fractions among samples.

We fit each pseudo-experiment using the same likelihood function used on data and we then derive the residual distributions (fitted value − input value). We choose to evaluate the residuals for the final observables that we are measuring, $R_{ADS}(h)$, $A_{ADS}(h)$ and $R^{\pm}(h)$. The residuals are fitted with a Gaussian function.

In principle, symptomatic behaviors can occur when the result of a parameter is close to its boundary. In the fit we impose the physical boundary that the fractions of signals must be positive. As can be seen in Table 5.1, none of the fractions is negative, but, the fraction of suppressed $B^- \to DK^-$ is very close to zero. And in fact this fit configuration provokes strong biases in the kaon observables.

In order to try to remove the biases we allow that fraction to vary in the range $[-1, 1]$. In principle we can also accept negative values for a fraction, if the final result is negatively small compared to the error.

Repeating the fit with the fraction of $B^- \to DK^-$ in $[-1, 1]$ range we obtain exactly the same results as in the previous case, but now there is no bias. For this reason we choose the final fit configuration where the $B^- \to DK^-$ fraction floats in the $[-1, 1]$ range.

In Figs. 5.8, 5.9, 5.10 and 5.11 we report the residuals (fitted values - input value) of the observables. The bias (the "p1" value in the legend of the figures) is computer
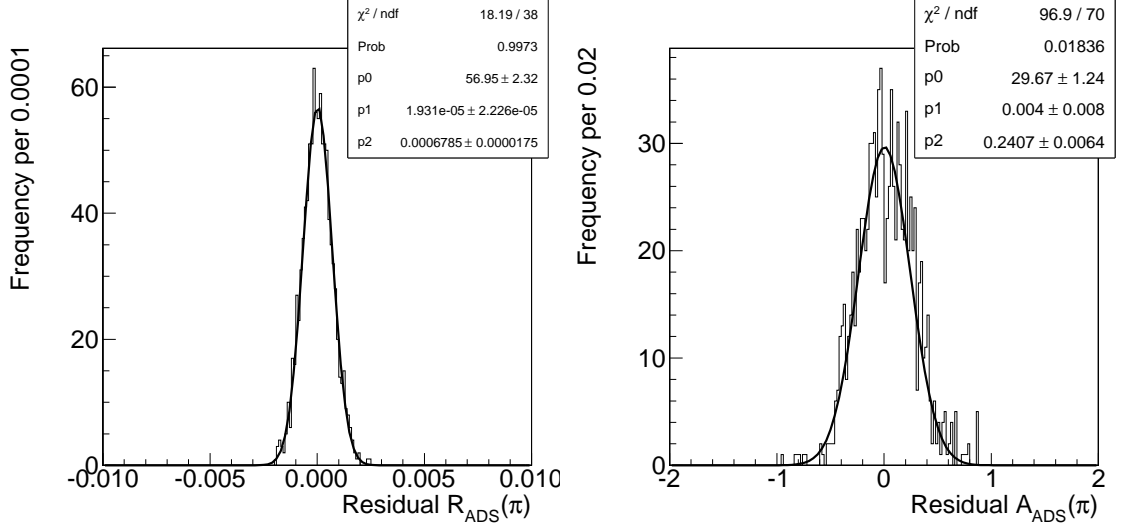
in fractions of sigmas (the "p2" value).

Figure 5.8: Residuals (fitted value - input value) of $R_{ADS}(\pi)$ (left) and $A_{ADS}(\pi)$ (right). Input values come from data. The distributions are fitted with a Gaussian function, where "p1" is the mean and "p2" is the sigma.
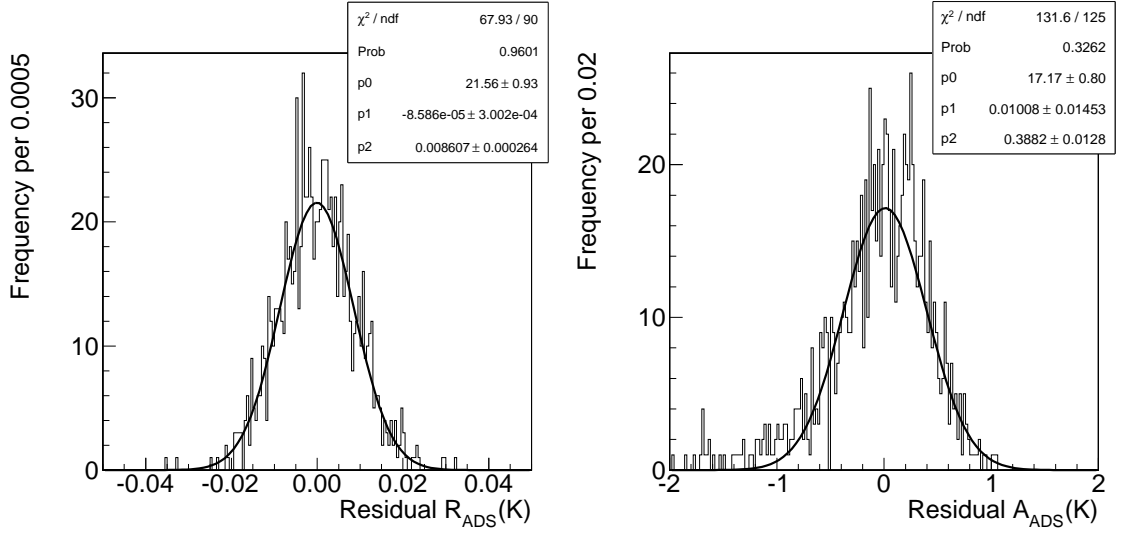

Figure 5.9: Residuals (fitted value - input value) of $R_{ADS}(K)$ (left) and $A_{ADS}(K)$ (right). Input values come from data. The distributions are fitted with a Gaussian function, where "p1" is the mean and "p2" is the sigma.

For $R_{ADS}(\pi)$ the bias is about $2 \cdot 10^{-5}$, which corresponds to about 0.03σ. For $A_{ADS}(\pi)$ the bias is about $4 \cdot 10^{-3}$, which corresponds to about 0.02σ. For $R_{ADS}(K)$
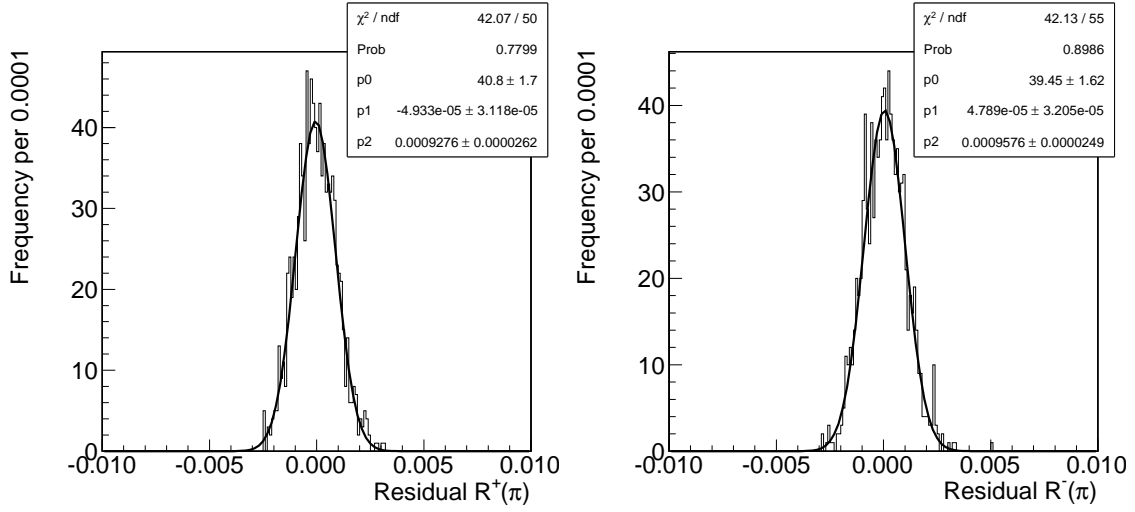
Figure 5.10: Residuals (fitted value - input value) of $R^+(\pi)$ (left) and $R^-(\pi)$ (right). Input values come from data. The distributions are fitted with a Gaussian function, where "p1" is the mean and "p2" is the sigma.
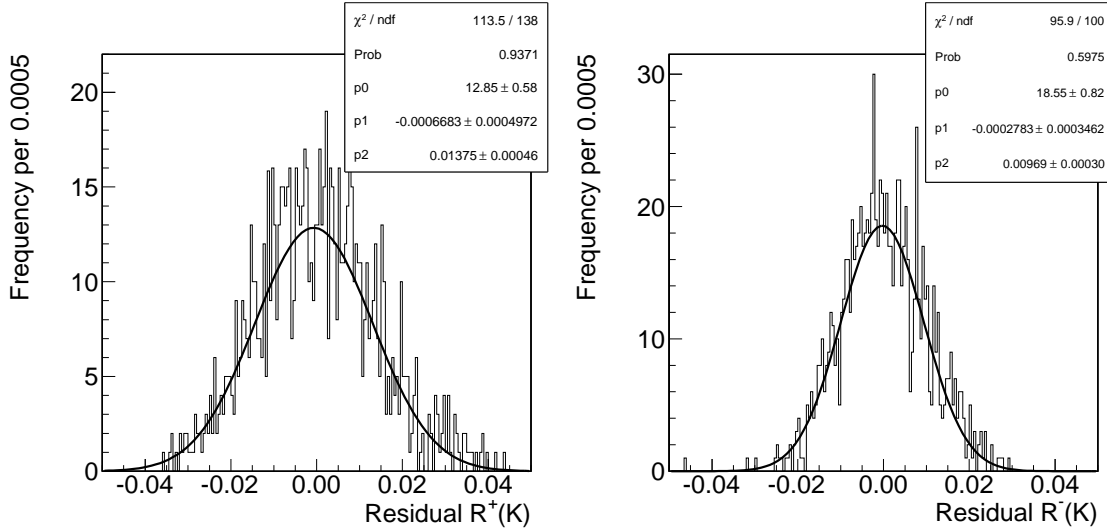

Figure 5.11: Residuals (fitted value - input value) of $R^+(K)$ (left) and $R^-(K)$ (right). Input values come from data. The distributions are fitted with a Gaussian function, where "p1" is the mean and "p2" is the sigma.

the bias is about $-9 \cdot 10^{-5}$, corresponding to about 0.01σ. For $A_{ADS}(K)$ the bias is about $1 \cdot 10^{-2}$, corresponding to about 0.03σ. For $R^+(\pi)$ the bias is about $-5 \cdot 10^{-5}$, which corresponds to about 0.05σ. For $R^-(\pi)$ the bias is about $5 \cdot 10^{-5}$, which corresponds to about 0.05σ. For $R^+(K)$ the bias is about $-7 \cdot 10^{-4}$, which

corresponds to about 0.05σ. For $R^-(K)$ the bias is about $3\cdot10^{-4}$, which corresponds to about 0.03σ.

From these values we can conclude that our fit does not have any significant bias.

5.4 Likelihood profile

In the proximity of the maximum $(\theta \approx \hat{\theta})$ a regular likelihood function can be approximated by a second-order Taylor expansion:

$$\log \mathcal{L}(\theta) \approx \log \mathcal{L}(\hat{\theta}) + \left[\frac{\partial \log \mathcal{L}}{\partial \theta}\right]_{\theta=\hat{\theta}} (\theta - \hat{\theta}) - \frac{1}{2}\left[\frac{\partial^2 \log \mathcal{L}}{\partial \theta^2}\right]_{\theta=\hat{\theta}} (\theta - \hat{\theta})^2 \qquad \boxed{5.3}$$

which can be re-written in the case of a multi-dimensional problem $(\vec{\theta})$ as:

$$\log \mathcal{L}(\vec{\theta}) \approx \log \mathcal{L}(\vec{\hat{\theta}}) - \frac{1}{2}\left[(\vec{\theta})^T V^{-1} \vec{\theta}\right], \qquad \boxed{5.4}$$

where $(\vec{\theta})^T$ is the transposed vector of the fit parameters and V^{-1} is the inverse of the covariance matrix. The expressions (5.3) and (5.4) imply that the contours of $\log \mathcal{L}$, locally around its maximum, can be approximated by an n-dimensional ellipsoid, where n is the multiplicity of the free parameters of the fit. The axes of the ellipsoid depend on the estimated correlation coefficients.

We evaluated the contour of the likelihood for the two more relevant variables, the positive and negative fraction of the suppressed $B^- \to DK^-$ decay, minimizing each time the likelihood with respect to the other variables ("Likelihood profile"). In Fig. 5.12 we can see the contours evaluated at a level of one (red), two (green) and three (blue) sigmas.

To obtain these distributions we use the `MNContour` function of `MINUIT` [74, 75], which plots one variable as a function of the other, taking into account the correlations between the two with a very accurate method, called `MINOS`, evaluating, at each point of the profile, the minimum of the likelihood function with respect to all the other $N-2$ parameters.

Small deviations from the expected ellipsoid shape are detected in the further contour of the three sigmas, which become more regular as we approach to the minimum. We still consider this behavior sufficiently regular for the purpose of our measurement, that is to first put into evidence the presence of these new signals and

make a first evaluation of the asymmetry.
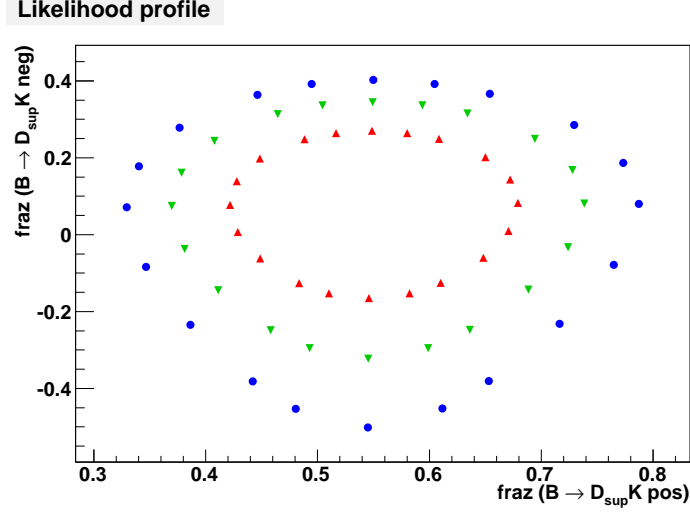

Figure 5.12: Likelihood profiles for the positive vs negative fractions of suppressed $B^- \to DK^-$, at the level of one (red), two (green) and three (blue) sigmas.

5.5 Correlation matrix

We study the correlation matrix (see Fig. 5.13) as returned by the fit, to search for possible large correlations. The coefficients of the matrix are defined as $\rho_{ij} = Cov(\hat{\theta}_i, \hat{\theta}_j)/\hat{\sigma}_{\hat{\theta}_i}\hat{\sigma}_{\hat{\theta}_j}$, where $Cov(\hat{\theta}_i, \hat{\theta}_j)$ is the off-diagonal element. The parameters of the matrix correspond to:

p1 fraction of positive favored background

p2 fraction of positive suppressed background

p3 fraction of negative favored background

p4 fraction of negative suppressed background

p5 fraction of positive favored $(B \to D\pi)$

p6 fraction of positive suppressed $(B \to DK)$

p7 fraction of negative favored $(B \to D\pi)$

p8 fraction of negative suppressed $(B \to DK)$

p9 constant of positive $(B \to D^*\pi)/(B \to D\pi)$

p10 constant of negative $(B \to D^*\pi)/(B \to D\pi)$

p11 fraction of positive $B \to [X]_D K$ background

p12 fraction of negative $B \to [X]_D K$ background

p13 fraction of positive $B \to [X]_D \pi$ background

p14 fraction of negative $B \to [X]_D \pi$ background

p15 fraction of positive $B \to K\pi\pi$ background

p16 fraction of negative $B \to K\pi\pi$ background

p17 fraction of positive $B^0 \to D^{*-}\ell\nu$ background

p18 fraction of negative $B^0 \to D^{*-}\ell\nu$ background

p21 slope of the combinatorial background

p23 fraction of π in the combinatorial background

p27 total number of positive favored events

p28 total number of negative favored events

p29 total number of positive suppressed events

p30 total number of negative suppressed events

Larger correlations come from the fractions of the background contributions in the suppressed sample. In particular between the fractions of $B \to [X]_D \pi$ and $B \to K\pi\pi$ (the two most contributing backgrounds in the suppressed sample), between the two and the total number of suppressed events, between the two and the total fraction of background. This effect is probably due to the fact that the fit is not really able to distinguish among those backgrounds, since they lies in almost the same mass region and have the same PID information.

Other correlations come from the fraction of $B^- \to D\pi^-$ and the ratio of $(B \to D^*\pi)/(B \to D\pi)$ (obviously correlated) and from the fraction of suppressed $B^- \to DK^-$ and the fraction of negative suppressed background. The latter is probably due to the fact that the fraction of $B^- \to DK^-$ is so small that can be confused with the background.

```
PARAMETER  CORRELATION COEFFICIENTS
  NO. GLOBAL    1     2     3     4     5     6     7     8     9    10    11    12    13    14    15    16
   1 0.33343  1.000 0.070 0.090 0.079 0.095 0.009 0.019-0.051-0.084-0.018-0.036-0.022-0.081-0.062-0.068-0.051
   2 0.91347  0.070 1.000 0.070 0.064 0.028 0.135 0.009-0.044-0.036 0.000-0.141-0.019-0.514-0.050-0.439-0.041
   3 0.32417  0.090 0.070 1.000 0.075 0.019 0.008 0.089-0.045-0.019-0.079-0.019-0.037-0.057-0.083-0.048-0.069
   4 0.91040  0.079 0.064 0.075 1.000 0.012 0.000 0.033-0.490 0.001-0.044-0.019-0.146-0.050-0.503-0.042-0.426
   5 0.50132  0.095 0.028 0.019 0.012 1.000 0.014 0.004-0.015-0.404-0.003-0.201-0.004 0.039-0.011 0.033-0.009
   6 0.17462  0.009 0.135 0.008 0.000 0.014 1.000-0.007 0.012-0.026 0.012-0.029-0.001-0.058-0.004-0.048-0.003
   7 0.51051  0.019 0.009 0.089 0.033 0.004-0.007 1.000-0.008-0.003-0.418-0.003-0.198-0.009 0.039-0.008 0.033
   8 0.50915 -0.051-0.044-0.045-0.490-0.015 0.012-0.008 1.000 0.010 0.003 0.012 0.055 0.031 0.246 0.026 0.211
   9 0.43929 -0.084-0.036-0.019 0.001-0.404-0.026-0.003 0.010 1.000 0.002-0.065 0.002 0.001 0.005 0.000 0.004
  10 0.45289 -0.018 0.000-0.079-0.044-0.003 0.012-0.418 0.003 0.002 1.000 0.002-0.065 0.005 0.002 0.005 0.001
  11 0.37500 -0.036-0.141-0.019-0.019-0.201-0.029-0.003 0.012-0.065 0.002 1.000 0.005 0.248 0.014 0.211 0.012
  12 0.38710 -0.022-0.019-0.037-0.146-0.004-0.001-0.198 0.055 0.002-0.065 0.005 1.000 0.014 0.263 0.012 0.221
  13 0.95391 -0.081-0.514-0.057-0.050 0.039-0.058-0.009 0.031 0.001 0.005 0.248 0.014 1.000 0.039 0.772 0.033
  14 0.95092 -0.062-0.050-0.083-0.503-0.011-0.004 0.039 0.246 0.005 0.002 0.014 0.263 0.039 1.000 0.033 0.760
  15 0.80898 -0.068-0.439-0.048-0.042 0.033-0.048-0.008 0.026 0.000 0.005 0.211 0.012 0.772 0.033 1.000 0.027
  16 0.79931 -0.051-0.041-0.069-0.426-0.009-0.003 0.033 0.211 0.004 0.001 0.012 0.221 0.033 0.760 0.027 1.000
  17 0.13371 -0.016-0.048-0.012-0.010 0.006-0.009-0.002 0.005 0.000 0.002 0.033 0.003 0.122 0.008 0.104 0.007
              1.000 0.002 0.040-0.011 0.021 0.000-0.114-0.002
  18 0.12978 -0.013-0.010-0.016-0.045-0.003-0.002 0.006 0.015 0.002 0.000 0.003 0.034 0.008 0.118 0.007 0.100
              0.002 1.000 0.042-0.007 0.000 0.020-0.002-0.111
  21 0.51431 -0.303-0.241-0.296-0.267-0.066-0.022-0.063 0.175 0.060 0.059 0.066 0.074 0.191 0.206 0.160 0.171
              0.040 0.042 1.000-0.020 0.007 0.007-0.039-0.041
  23 0.19892  0.021-0.005 0.032-0.047-0.014 0.104-0.013 0.147-0.027-0.026 0.004 0.013-0.022 0.005-0.017 0.006
             -0.011-0.007-0.020 1.000-0.005-0.004 0.025 0.023
  27 0.47222 -0.002 0.022-0.002-0.002 0.000 0.005-0.001 0.001 0.000 0.001 0.039 0.000 0.131 0.001 0.111 0.001
              0.021 0.000 0.007-0.005 1.000 0.000 0.003 0.000
  28 0.45740 -0.002-0.002-0.002 0.021-0.001-0.001 0.000-0.010 0.001 0.000 0.000 0.041 0.001 0.131 0.001 0.109
              0.000 0.020 0.007-0.004 0.000 1.000 0.000 0.003
  29 0.94517  0.011-0.118 0.012 0.009 0.002-0.027 0.003-0.003 0.000-0.006-0.212-0.002-0.721-0.008-0.609-0.006
             -0.114-0.002-0.039 0.025 0.003 0.000 1.000 0.002
  30 0.94253  0.013 0.010 0.012-0.114 0.003 0.003 0.002 0.058-0.006 0.001-0.002-0.228-0.008-0.728-0.007-0.610
             -0.002-0.111-0.041 0.023 0.000 0.003 0.002 1.000
```

Figure 5.13: Correlation matrix of the fitted parameters.

5.6 Efficiency corrections

From the yields obtained in Table 5.1 we can immediately compute raw values of the observables, according to:

$$
\begin{aligned}
R_{ADS}(h) &= \frac{BR(B^- \to D(\to K^+\pi^-)h^-) + BR(B^+ \to D(\to K^-\pi^+)h^+)}{BR(B^- \to D(\to K^-\pi^+)h^-) + BR(B^+ \to D(\to K^+\pi^-)h^+)} \\
A_{ADS}(h) &= \frac{BR(B^- \to D(\to K^+\pi^-)h^-) - BR(B^+ \to D(\to K^-\pi^+)h^+)}{BR(B^- \to D(\to K^+\pi^-)h^-) + BR(B^+ \to D(\to K^-\pi^+)h^+)} \quad (5.5) \\
R^{\pm}(h) &= \frac{BR(B^{\pm} \to D(\to K^{\mp}\pi^{\pm})h^{\pm})}{BR(B^{\pm} \to D(\to K^{\pm}\pi^{\mp})h^{\pm})}
\end{aligned}
$$

Errors are evaluated with the formula in eq. (5.2).

The results are:

$$
\begin{aligned}
R_{ADS}(\pi)_{raw} &= (2.8 \pm 0.7 \ (stat.)) \cdot 10^{-3} \\
R_{ADS}(K)_{raw} &= (22.2 \pm 8.6 \ (stat.)) \cdot 10^{-3} \\
A_{ADS}(\pi)_{raw} &= 0.13 \pm 0.25 \ (stat.) \\
A_{ADS}(K)_{raw} &= -0.82 \pm 0.44 \ (stat.) \\
R^+(\pi)_{raw} &= (2.4 \pm 1.0 \ (stat.)) \cdot 10^{-3} \quad (5.6) \\
R^-(\pi)_{raw} &= (3.1 \pm 1.1 \ (stat.)) \cdot 10^{-3} \\
R^+(K)_{raw} &= (42.6 \pm 13.7 \ (stat.)) \cdot 10^{-3} \\
R^-(K)_{raw} &= (3.8 \pm 10.3 \ (stat.)) \cdot 10^{-3}
\end{aligned}
$$

where the subscript "raw" means that the values are not yet been corrected for the relative efficiencies.

In fact, in order to traslate the raw result of the fit into measurement of relative BRs and CP-violating asymmetries we need to apply corrections for the different relative efficiencies of positively and negatively charged particles interacting in the detectors material, because different hadronic interaction between K^+ and K^- [2] (π^+ and π^-) can fake the measurement. This effect is reproduced rather well by Monte Carlo simulations and we use the results of [77] for the K^+/K^- efficiency:

$$
\frac{\epsilon(K^+)}{\epsilon(K^-)} = 1.0178 \pm 0.0023(\text{stat.}) \pm 0.0045(\text{syst.}) \quad (5.7)
$$

---

[2] K^- has a larger hadronic cross section than K^+.

and for the π^+/π^- efficiency:

$$\frac{\epsilon(\pi^+)}{\epsilon(\pi^-)} = 0.997 \pm 0.003(\text{stat.}) \pm 0.006(\text{syst.}). \tag{5.8}$$

The efficiency of kaon and pion from D is instead evaluated on our favored $B^- \to D\pi^-$ sample and is equal to:

$$\frac{\epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-)} = 0.998 \pm 0.015(\text{stat}) \pm 0.016(\text{syst}). \tag{5.9}$$

More details on how we extract this value can be found in the following sections.

5.7 $\frac{K^-\pi^+}{K^+\pi^-}$ efficiency

We can evaluate the $K^-\pi^+/K^+\pi^-$ efficiency directly on favored sample of $B^- \to D\pi^-$ decay, for which we expect zero asymmetry. The value from [16], which accounts only for a Belle result of 2006, is -0.008 ± 0.008, compatible with zero. On our data we found $N(B^+) = 9881 \pm 103$ and $N(B^-) = 9893 \pm 103$, so the asymmetry of the favored pion mode[3] is equal to $A^m_{fav}(\pi) = (0.6 \pm 7.3) \cdot 10^{-3}$.

To evaluate the efficiency we can proceed in this way.

5.7.1 Strategy

- The central value of the efficiency is evaluated considering a true value of the asymmetry equal to zero.

- The statistical error is obtained propagating the error on the efficiency of the pion from B and the error on the measured asymmetry.

- The systematic error is obtained propagating the error on the pdg measurement of the asymmetry.

---

[3]The asymmetry is calculated using all digits of the numbers of events.

5.7.2 Calculation

How to compare the measured asymmetry with the true asymmetry and the efficiency

The number of measured events (N^m) can be written as $N^t \cdot \epsilon$, where N^t is the true number of events and ϵ is the efficiency of reconstruction.

The true value of the asymmetry is:

$$A^t_{fav}(\pi) = \frac{N([K^-\pi^+]\pi^-) - N([K^+\pi^-]\pi^+)}{N([K^-\pi^+]\pi^-) + N([K^+\pi^-]\pi^+)} = \frac{N^{t-}_f - N^{t+}_f}{N^{t-}_f + N^{t+}_f}$$

$$= \frac{N^{t-}_f/N^{t+}_f - 1}{N^{t-}_f/N^{t+}_f + 1}$$

from which $N^{t-}_f/N^{t+}_f = \dfrac{1 + A^t_{fav}}{1 - A^t_{fav}}$.

The total efficiency (ϵ) we are considering is:

$$\epsilon^- = \epsilon(K^-\pi^+) \cdot \epsilon(\pi^-) = \epsilon^-_D \cdot \epsilon^-_\pi$$

$$\epsilon^+ = \epsilon(K^+\pi^-) \cdot \epsilon(\pi^+) = \epsilon^+_D \cdot \epsilon^+_\pi.$$

so we can write the measured asymmetry:

$$A^m_{fav}(\pi) = \frac{N^{m-}_f - N^{m+}_f}{N^{m-}_f + N^{m+}_f} = \frac{N^{t-}_f \cdot \epsilon^- - N^{t+}_f \cdot \epsilon^+}{N^{t-}_f \cdot \epsilon^- + N^{t+}_f \cdot \epsilon^+}$$

$$= \frac{N^{t-}_f/N^{t+}_f \cdot \epsilon^-/\epsilon^+ - 1}{N^{t-}_f/N^{t+}_f \cdot \epsilon^-/\epsilon^+ + 1}$$

from which $\epsilon^-/\epsilon^+ = \dfrac{1}{N^{t-}_f/N^{t+}_f} \cdot \dfrac{1 + A^m}{1 - A^m}$,

$$\frac{\epsilon^-_D}{\epsilon^+_D} \cdot \frac{\epsilon^-_\pi}{\epsilon^+_\pi} = \left(\frac{1 - A^t}{1 + A^t}\right)\left(\frac{1 + A^m}{1 - A^m}\right)$$

$$\downarrow$$

$$\frac{\epsilon^-_D}{\epsilon^+_D} = \frac{\epsilon^+_\pi}{\epsilon^-_\pi} \cdot \left(\frac{1 - A^t}{1 + A^t}\right)\left(\frac{1 + A^m}{1 - A^m}\right) \qquad \boxed{5.10}$$

5. Extraction of the results

Efficiency central value

To evaluate the central value we use $A^t = 0$,

$$\frac{\epsilon_D^-}{\epsilon_D^+} = \frac{\epsilon_\pi^+}{\epsilon_\pi^-} \cdot \left(\frac{1+A^m}{1-A^m}\right)$$

(5.11)

and, since $\epsilon_\pi^+/\epsilon_\pi^- = 0.997 \pm 0.003 \pm 0.006$:

$$\frac{\epsilon_D^-}{\epsilon_D^+} = \frac{\epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-)} = 0.998.$$

Efficiency statistical error

Propagating the error on the pion efficiency and on the measured asymmetry from eq. (5.11) we obtain this expression:

$$\Delta\frac{\epsilon_D^-}{\epsilon_D^+}(\text{stat}) = \sqrt{\left(\frac{1+A^m}{1-A^m}\right)^2 (\Delta\epsilon_\pi^-/\epsilon_\pi^+)^2 + \left(\frac{2\epsilon_\pi^-/\epsilon_\pi^+}{(1-A^m)^2}\right)^2 (\Delta A^m)^2} = 0.015.$$

Efficiency systematic error

To evaluate the systematic error we can use the complete formula (5.10) and propagate the error from the value listed in [16].

$$\Delta\frac{\epsilon_D^-}{\epsilon_D^+}(syst) = \left(\frac{1+A^m}{1-A^m}\right) \cdot \frac{2\epsilon_\pi^-/\epsilon_\pi^+}{(1+A^t)^2} \cdot (\Delta A^t) = 0.016.$$

The final result is:

$$\frac{\epsilon_D^-}{\epsilon_D^+} = \frac{\epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-)} = 0.998 \pm 0.015(stat) \pm 0.016(syst).$$

We apply the corrections to the observables, according to these formulas:

$$R_{ADS}(h) = \frac{N_s^- \epsilon(\frac{K^-\pi^+}{K^+\pi^-})\epsilon(\frac{h^+}{h^-}) + N_s^+}{N_f^- \epsilon(\frac{K^+\pi^-}{K^-\pi^+})\epsilon(\frac{h^+}{h^-}) + N_f^+}$$

$$A_{ADS}(h) = \frac{N_s^- \epsilon(\frac{K^-\pi^+}{K^+\pi^-})\epsilon(\frac{h^+}{h^-}) - N_s^+}{N_s^- \epsilon(\frac{K^-\pi^+}{K^+\pi^-})\epsilon(\frac{h^+}{h^-}) + N_s^+}$$

$$R^\pm(h) = \frac{N_s^\pm}{N_f^\pm}\epsilon(\frac{K^\pm\pi^\mp}{K^\mp\pi^\pm})$$

where $N_s^\pm$ are the number of positive and negative suppressed events, $N_f^\pm$ are the number of positive and negative favored events and h is pion or kaon.

Final corrected results are:

$$
\begin{aligned}
R_{ADS}(\pi) &= (2.8 \pm 0.7) \cdot 10^{-3} \\
R_{ADS}(K) &= (22.0 \pm 8.6) \cdot 10^{-3} \\
A_{ADS}(\pi) &= 0.13 \pm 0.25 \\
A_{ADS}(K) &= -0.82 \pm 0.44 \\
R^+(\pi) &= (2.4 \pm 1.0) \cdot 10^{-3} \\
R^-(\pi) &= (3.1 \pm 1.1) \cdot 10^{-3} \\
R^+(K) &= (42.6 \pm 13.7) \cdot 10^{-3} \\
R^-(K) &= (3.8 \pm 10.3) \cdot 10^{-3}
\end{aligned}
\tag{5.12}
$$

There are almost no changes from the uncorrected results, the only difference being on $R_{ADS}(K)$. Anyway the variation is only a small fraction of the statistical error.

The asymmetry of the pion mode is compatible with zero, while the asymmetry of the kaon mode is close to -1 and about 2σ away from zero. It is interesting to exactly evaluate how significantly is this value different from zero.

5.8 Significance of the suppressed $B^- \to DK^-$ asymmetry

To evaluate the significance of $A_{ADS}(K)$ we generate 2000 pseudo-experiments with zero input value of the asymmetry and the other values as found on data. Fitting these pseudo-experiment with our central fit, we obtain the distribution of the asymmetry as in Fig. 5.14.

The distribution is fitted with a gaussian distribution with fixed mean equal to zero. The sigma of this distribution is 0.37 and, with this value, our central value of the asymmetry is 2.2σ away from zero. We cannot therefore establish an evidence for the presence of such asymmetry, in spite of the suggestive appearance of the plots.

Using the equations (1.5.1), the values listed in Sect. 1.5.1, and the average of γ from [21], we can give an estimate of the expected value for $A_{ADS}(K)$, which is about -0.62. We find our measurement in agreement with the expectation.
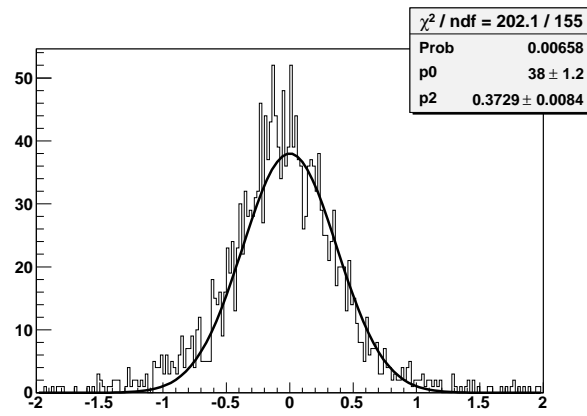
Figure 5.14: Distribution of the kaon asymmetry from 2000 fit on pseudo-experiments with zero input value of the asymmetry. The distribution is fitted with a gaussian distribution with fixed mean equal to zero.

6

Systematic uncertainties

In this chapter we will describe the contributions to the systematic uncertainty of the measurement coming from the mass and dE/dx models, and from the efficiency. We will describe also the evaluation of the significance of the suppressed $B^- \to DK^-$ signal including also the systematic effects.

6.1 Systematic uncertainties

6.1.1 PID model

As explained in Sect. 4.6 the PID information is contained in the kaoness variable κ, whose distributions for pions and kaons are shown in Fig. 4.18. Both distributions are taken from data and they are well parameterized with the sum of three gaussian.

The general procedure followed in previous analysis in CDF for the systematic evaluation, as in [78] or [79], was to repeat the fit randomly varying the parameters of the Likelihood functions independently within a 1σ radius multi-dimensional sphere in the space of the parameters of the dE/dx calibrations.

This method has been improved in the current Thesis considering also the correlations among those parameters.

The detailed procedure is described in the follows.

The nine parameters of the gaussians modeling the PID distributions are correlated according to a specific correlation matrix found on data. From these nine

parameters and their correlation matrix we can generate new sets of parameters, randomly varying them according to a multi-variate gaussian distributions with mean values set to those found on data and the same correlation matrix.

In order to numerically generate multivariate gaussian variables for our simulation, we have used the following algorithm [16].

We start from a $n \times n$ covariance matrix V, with components $V_{ij} = \rho_{ij}\sigma_i\sigma_j$. We first generate n independent gaussian variables $\{\eta_j\}$, with mean 0 and variance 1. The new set of variables $\{x_i\}$ is obtained as $x_i = \mu_i + \sum_j L_{ij}\eta_j$, where μ_i is the mean of x_i, and L_{ij} are the components of L, the unique lower triangular matrix that fulfills $V = LL^T$.

The matrix L is computed using a recursive relation (Cholesky method):

$$
\begin{aligned}
L_{jj} &= \left(V_{jj} - \sum_{k=1}^{j-1} L_{jk}^2 \right)^{1/2} \\
L_{ij} &= \frac{V_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{jk}}{L_{jj}} \qquad \text{for } j = 1, \ldots, n \text{ and } i = j+1, \ldots, n.
\end{aligned}
\tag{6.1}
$$

The matrix L is computed at the beginning of the iteration and then, after generating the set of $\{\eta_j\}$, we can easily get the $\{x_i\}$ set.

For each set of $\{x_i\}$ we have a different κ curve. We generate one thousand of them, shown in Fig. 6.1 (red lines) superimposed to the data (points). The curves completely cover the region bound by error data, as expected.

Each systematic curve is used to fit the data. In this way we obtain one thousand values for each observable. We draw the histograms of the observables and, fitting them with gaussian functions, we take the sigma of the gaussian as the dE/dx systematic error.

In Figs. 6.2, 6.3, 6.4 and 6.5 we can see respectively the distributions of $A_{ADS}(K)$, $A_{ADS}(\pi)$, $R_{ADS}(K)$, $A_{ADS}(\pi)$, $R^{\pm}(K)$ and $R^{\pm}(\pi)$ obtained from the one thousand of fits on data. The "p2" value in the legend is the sigma of the gaussian, which is taken as systematic error.
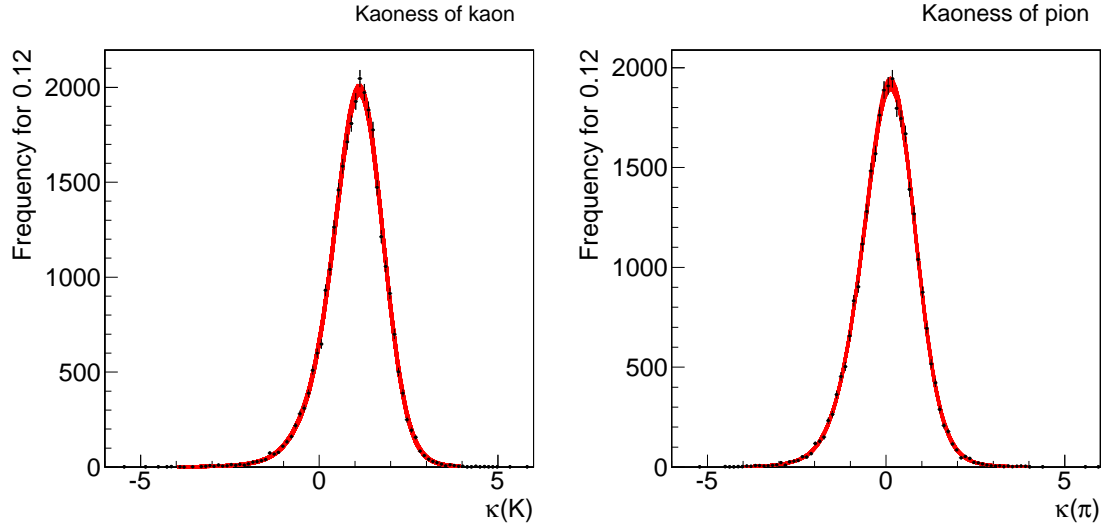
Figure 6.1: κ variable, on the left for the K, on the right for the π. Points with error bars come from data, in red are one thousand systematic curves.
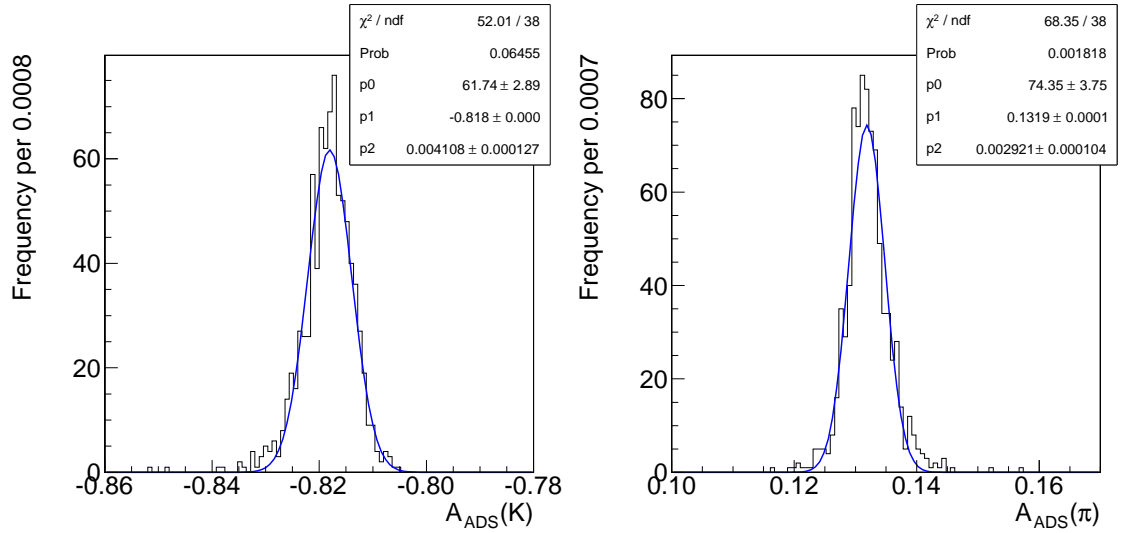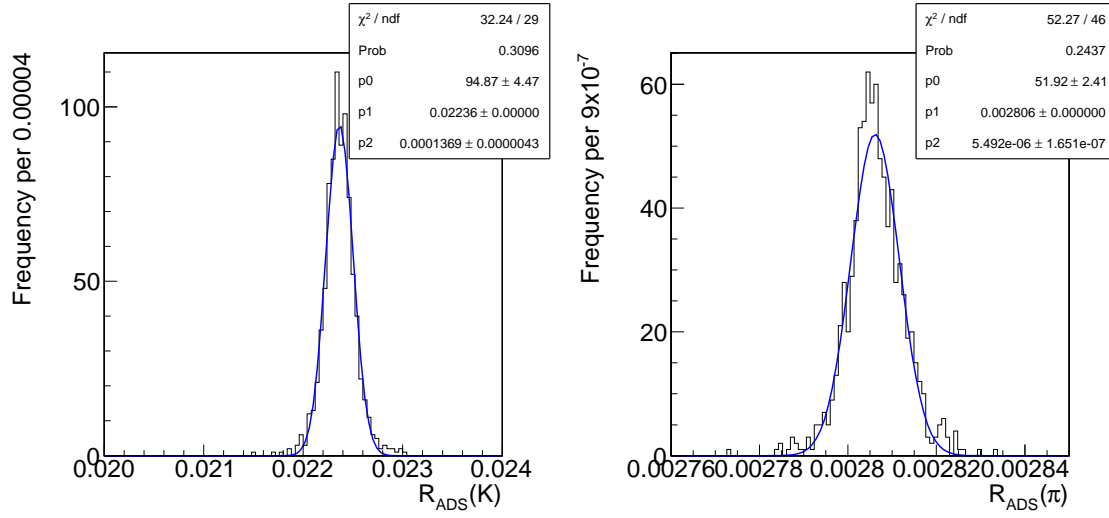

Figure 6.2: Distributions of $A_{ADS}(K)$ (left) and $A_{ADS}(\pi)$ (right) obtained with the systematic curves of dE/dx. They are fitted with a gaussian and the sigma ("p2") is taken as systematic error.
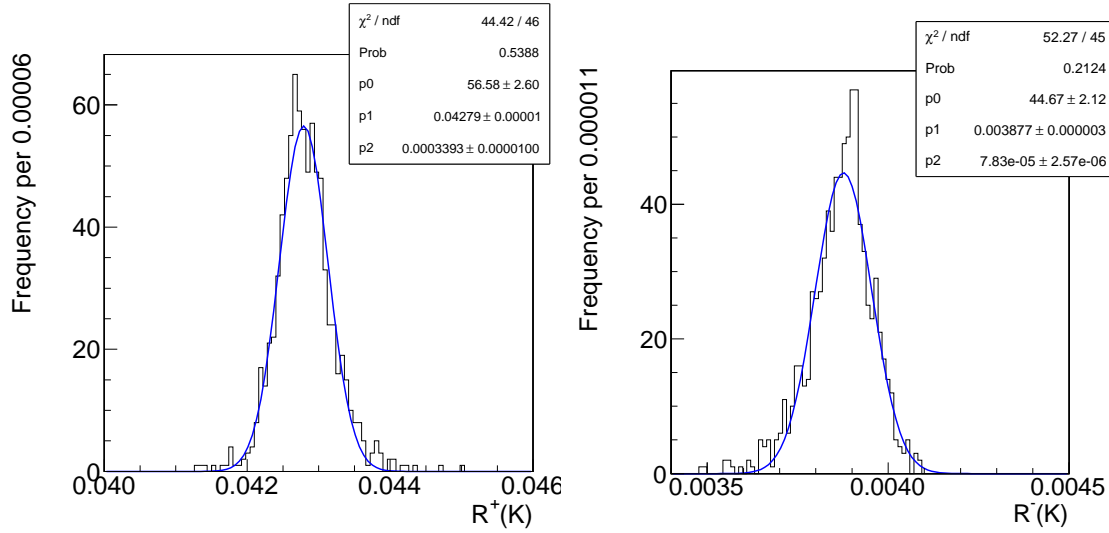
Figure 6.3: Distributions of $R_{ADS}(K)$ (left) and $R_{ADS}(\pi)$ (right) obtained with the systematic curves of dE/dx. They are fitted with a gaussian and the sigma ("p2") is taken as systematic error.
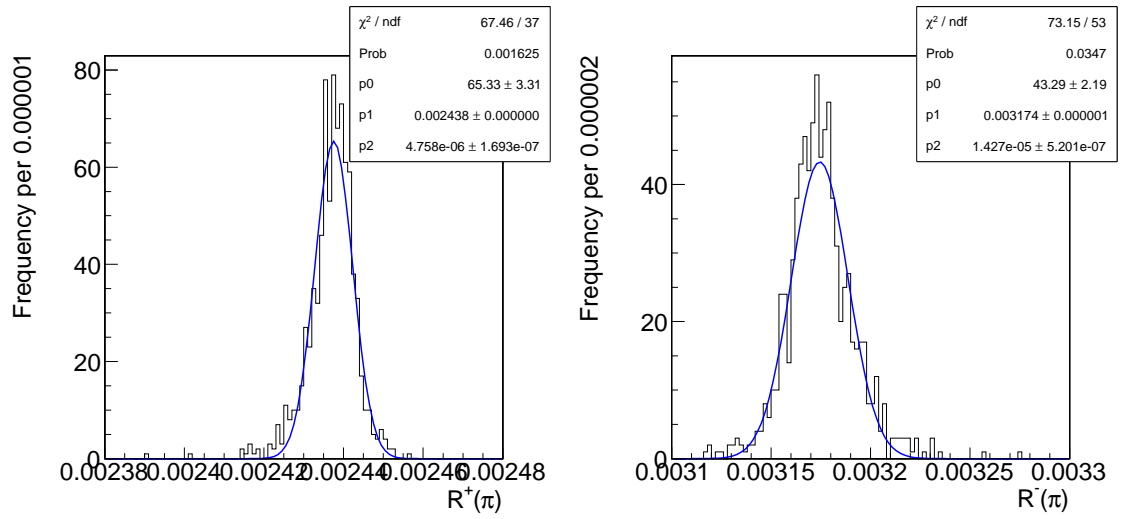

Figure 6.4: Distributions of $R^{+}(K)$ (left) and $R^{-}(K)$ (right) obtained with the systematic curves of dE/dx. They are fitted with a gaussian and the sigma ("p2") is taken as systematic error.

Figure 6.5: Distributions of $R^+(\pi)$ (left) and $R^-(\pi)$ (right) obtained with the systematic curves of dE/dx. They are fitted with a gaussian and the sigma ("p2") is taken as systematic error.

6.1.2 Mass model of the combinatorial background

In the central fit we assume that the exponential function modeling the combinatorial background is valid also under the peak of the B signals. This has been verified fitting two different samples made up only by combinatorial events:

- B events from the D sidebands (outside 3σ from the central D mass value) (Fig. 6.6);

- B events selected with the same criteria of this analysis, but $\chi^2_{3D} \geq 30$.

They are two samples independent from each other and also independent from the samples used in the analysis. In Fig. 6.6 we can see the two distributions fitted with an exponential function. Both of them are properly modeled by that function in the whole range.
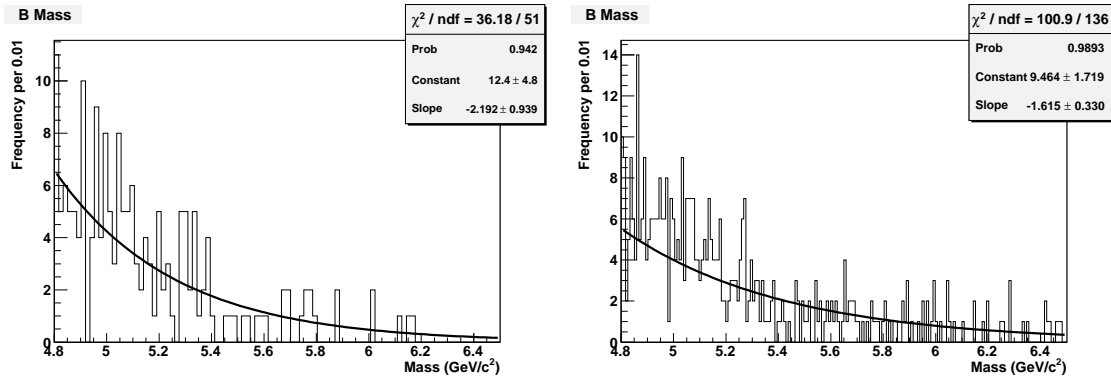

Figure 6.6: Invariant B mass distribution in the D sidebands (left) and with the $\chi^2_{3D} \geq 30$ (right) fitted with an exponential distribution. The exponential fit properly models the events.

To evaluate the systematic effect due to the combinatorial background shape, we redo the fit on data using different shapes, each of them a variation of the exponential function:

- exponential plus constant;

- exponential plus first degree polynomial;

- exponential plus second degree polynomial.

We then take as systematic error the worse result of the three cases, corresponding to the case when the fitted value of the slope of exponential is the most distant from the central value. It occurs when we use the exponential plus constant, where the slope goes from -2.61 ± 0.13 of the central fit, to -2.64 ± 0.24. In the other two cases the value of the slope remain the same.

We used as systematic error the difference between the observables results obtained using the exponential plus constant and the central fit.

6.1.3 Physics background mass model

To evaluate the contribution from the physics background mass model, we varied the shapes in the following ways:

- $B^- \to D(\to X)\pi^-$

  The mass shape is parametrized with a Pearson function of the IV type, we vary the parameters corresponding to the mean and the width of $\pm 1\sigma$.

- $B^- \to D(\to X)K^-$

  Since we parametrized the mass shape with a gaussian, we vary the mean and the width of the gaussian of $\pm 1\sigma$.

- $B^- \to K^-\pi^+\pi^-$

  We parametrized the mass shape with the sum of two gaussians. To evaluate the systematics we vary the mean and the width of the second gaussian, that is the one going under the B signal peaks, of $\pm 1\sigma$.

- $B^0 \to D_0^{*-}\ell^+\nu_\ell$

  We parametrized the mass shape with the sum of two gaussians. To evaluate the systematics we vary the mean and the width of the second gaussian, that is the one going under the B signal peaks, of $\pm 1\sigma$.

- $B^- \to D^{*0}\pi^-$

  We parametrized the mass shape with the sum of three gaussians plus an exponential. To evaluate the systematics we vary the slope of the exponential of $\pm 1\sigma$.

6.1.4 Efficiency

From eqs. (5.12) we can see the effect of the efficiencies $\frac{\epsilon(\pi^+)}{\epsilon(\pi^-)}$, $\frac{\epsilon(K^+)}{\epsilon(K^-)}$ and $\frac{\epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-)}$ in the observables. We propagate the error on the efficiencies to the observables using the standard formula for the error propagation.

The error on the efficiency is taken as the sum in quadrature of the statistical and systematic error of eqs. (5.7), (5.8) and (5.9). We will call it $\Delta\epsilon(i)$, where i is the type of efficiency.

An other way to propagate the error is to evaluate $\epsilon^-(i)$ and $\epsilon^+(i)$, given by $\epsilon^\pm(i) = \epsilon(i) \pm \Delta\epsilon(i)$, and substitute them in the eqs. (5.12). We get two values for each observable and we can take as systematic error the difference divided by two.

We find that the results in these two ways are the same.

6.1.5 Total systematic uncertainties

A summary of all systematics is reported in Table 6.1.

The total systematic uncertainty on each observable is the sum in quadrature of all systematic uncertainties. For completeness we report all digits in the Table, while in the final result we will drop the last digit.

Source	$R^+(\pi)$	$R^-(\pi)$	$R^+(K)$	$R^-(K)$	$R_{ADS}(\pi)$	$R_{ADS}(K)$	$A_{ADS}(\pi)$	$A_{ADS}(K)$
dE/dx	0.00001	0.00001	0.0003	0.0001	0.00001	0.0001	0.003	0.004
combinatorial background	0.00001	0.00001	0.0001	0.0002	0.00001	0.0001	0.001	0.010
$B^- \to [X]_D \pi^-$ shape	0.00040	0.00040	0.0026	0.0026	0.00038	0.0025	0.013	0.089
$B^- \to [X]_D K^-$ shape	0.00001	0.00001	0.0001	0.0001	0.00001	0.0001	0.000	0.003
$B^- \to K^-\pi^+\pi^-$ shape	0.00002	0.00002	0.0001	0.0001	0.00002	0.0001	0.001	0.003
$B^0 \to D_0^{*-}\ell^+\nu_\ell$ shape	0.00004	0.00004	0.0003	0.0002	0.00004	0.0002	0.003	0.007
$B^- \to D^{*0}\pi^-$ shape	0.00004	0.00005	0.0005	0.0004	0.00005	0.0005	0.001	0.014
efficiency of K_B	-	-	-	-	-	0.0001	-	0.002
efficiency of π_B	-	-	-	-	-	-	0.003	-
efficiency of $K_D \pi_D$	0.00005	0.00007	0.0009	0.0001	0.00006	0.0003	0.011	0.004
Total	**0.00041**	**0.00041**	**0.0028**	**0.0027**	**0.00039**	**0.0026**	**0.018**	**0.091**

Table 6.1: Systematic uncertainties for all observables.

6.2 Significance of suppressed $B^- \to DK^-$ signal

In Sect. 5.1 we evaluated the statistical significance of the suppressed $B^- \to DK^-$ signal. To include also the systematic variations we use a conservative procedure, called seupremum p-value.

We first need to find the worst-case systematic configuration, that is the one in which the DLL value is closest to zero. To do this we repeat the fit on data, using each time one different systematic shape. The only three systematics that have an effect of reducing the DLL are: the dE/dx, the $B^- \to D(\to X)\pi^-$ and the $B^0 \to D_0^{*-}\ell^+\nu_\ell$ shape.

We repeat again the fit on data combining in every possible way those three systematics and, for each fit we calculate the DLL, obtaining the distribution in Fig. 6.7. We select the configuration giving the lower DLL as the configuration for the significance evaluation.
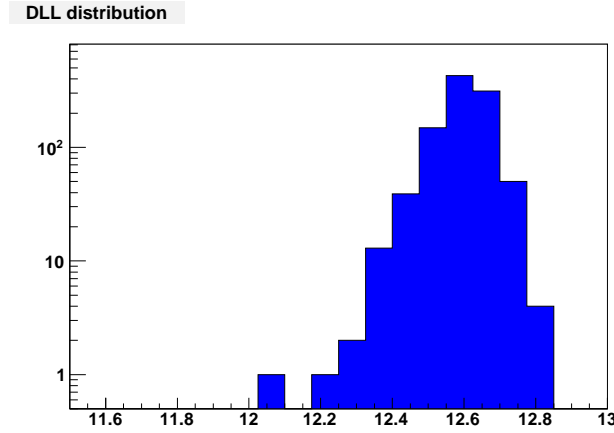

Figure 6.7: $-2\Delta \log \mathcal{L}$ distribution from fit on data, obtained combining various systematic configurations on dE/dx and physics background *pdf*s.

We then generate 48000 pseudo-experiments with zero $B^- \to DK^-$ signal events. Moreover the events generated have as dE/dx and physics background *pdf*s the configurations chosen in the previous step.

We fit these pseudo-experiments with the central fit, obtaining the distribution of DLL shown in Fig. 6.8. The number of cases with DLL greater than 14.1 (the

value found on data) is 69, which, divided by the total 46664, corresponds to a p-value of about $1.48 \cdot 10^{-3}$. The resulting significance including systematics is 3.2 σ, resulting in an evidence of the $B^- \to DK^-$ signal.
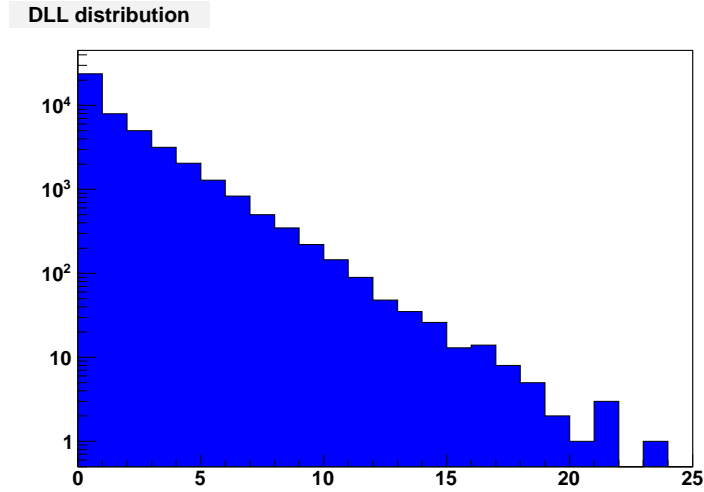

Figure 6.8: $-2\Delta \log \mathcal{L}$ distribution from the central fit on pseudo-experiment with zero $B^- \to DK^-$ signal input.

7
Final results and conclusions

Using 7 fb^{-1} of data collected by the CDF experiment we measure:

$$R_{ADS}(\pi) = \frac{BR(B^- \to D(\to K^+\pi^-)\pi^-) + BR(B^+ \to D(\to K^-\pi^+)\pi^+)}{BR(B^- \to D(\to K^-\pi^+)\pi^-) + BR(B^+ \to D(\to K^+\pi^-)\pi^+)} =$$

$$= (2.8 \pm 0.7 \ (stat.) \pm 0.4 \ (syst.)) \cdot 10^{-3}$$

$$R_{ADS}(K) = \frac{BR(B^- \to D(\to K^+\pi^-)K^-) + BR(B^+ \to D(\to K^-\pi^+)K^+)}{BR(B^- \to D(\to K^-\pi^+)K^-) + BR(B^+ \to D(\to K^+\pi^-)K^+)} =$$

$$= (22.0 \pm 8.6 \ (stat.) \pm 2.6 \ (syst.)) \cdot 10^{-3}$$

$$A_{ADS}(\pi) = \frac{BR(B^- \to D(\to K^+\pi^-)\pi^-) - BR(B^+ \to D(\to K^-\pi^+)\pi^+)}{BR(B^- \to D(\to K^+\pi^-)\pi^-) + BR(B^+ \to D(\to K^-\pi^+)\pi^+)} =$$

$$= 0.13 \pm 0.25 \ (stat.) \pm 0.02 \ (syst.)$$

$$A_{ADS}(K) = \frac{BR(B^- \to D(\to K^+\pi^-)K^-) - BR(B^+ \to D(\to K^-\pi^+)K^+)}{BR(B^- \to D(\to K^+\pi^-)K^-) + BR(B^+ \to D(\to K^-\pi^+)K^+)} =$$

$$= -0.82 \pm 0.44 \ (stat.) \pm 0.09 \ (syst.)$$

$$R^+(\pi) \;=\; \frac{BR(B^+ \to D(\to K^- \pi^+)\pi^+)}{BR(B^+ \to D(\to K^+ \pi^-)\pi^+)} =$$

$$=\; (2.4 \pm 1.0 \;(stat.) \pm 0.4 \;(syst.)) \cdot 10^{-3}$$

$$R^-(\pi) \;=\; \frac{BR(B^- \to D(\to K^+ \pi^-)\pi^-)}{BR(B^- \to D(\to K^- \pi^+)\pi^-)} =$$

$$=\; (3.1 \pm 1.1 \;(stat.) \pm 0.4 \;(syst.)) \cdot 10^{-3}$$

$$R^+(K) \;=\; \frac{BR(B^+ \to D(\to K^- \pi^+)K^+)}{BR(B^+ \to D(\to K^+ \pi^-)K^+)} =$$

$$=\; (42.6 \pm 13.7 \;(stat.) \pm 2.8 \;(syst.)) \cdot 10^{-3}$$

$$R^-(K) \;=\; \frac{BR(B^- \to D(\to K^+ \pi^-)K^-)}{BR(B^- \to D(\to K^- \pi^+)K^-)} =$$

$$=\; (3.8 \pm 10.3 \;(stat.) \pm 2.7 \;(syst.)) \cdot 10^{-3}$$

We find evidence of the suppressed $B^- \to DK^-$ and $B^- \to D\pi^-$ signals, respectively with a significance of 3.2σ (including systematics) and 3.6σ (only statistical).

This is the first time measurement of these quantities are done in a hadron collider environment. The results have been presented for the first time at the international conference CKM 2010 on September 2010, simultaneously with a similar measurement by Belle [29]. The results are in agreement and with comparable resolutions with respect to measurements from Belle and *BABAR* [28] (Figs. 7.1 and 7.2). The average has been done by the HFAG group [21]. Recently the LHCb experiment performed the same measurement, confirming these results with a better precision [30] (Figs. 7.3 and 7.4).

These results have already been combined together for the extraction of the γ angle [21].
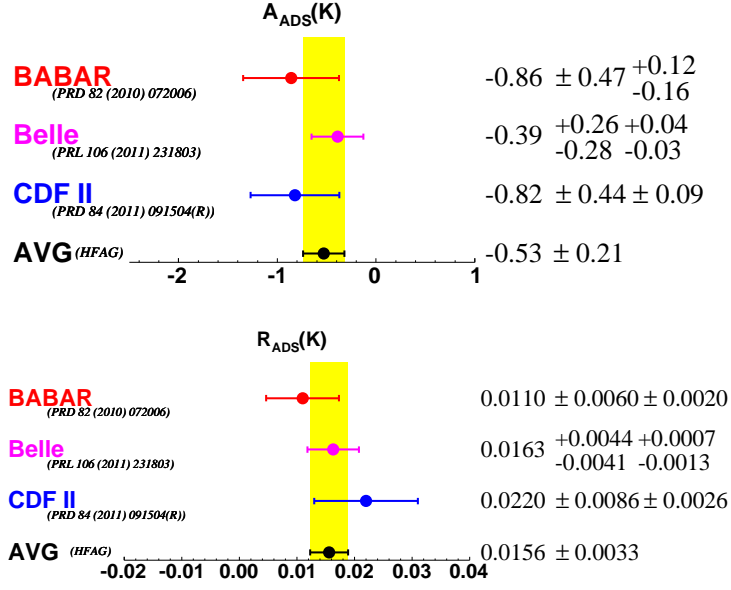
Figure 7.1: Comparison of $BABAR$, Belle and CDF experimental results on $A_{ADS}(K)$ and $R_{ADS}(K)$.
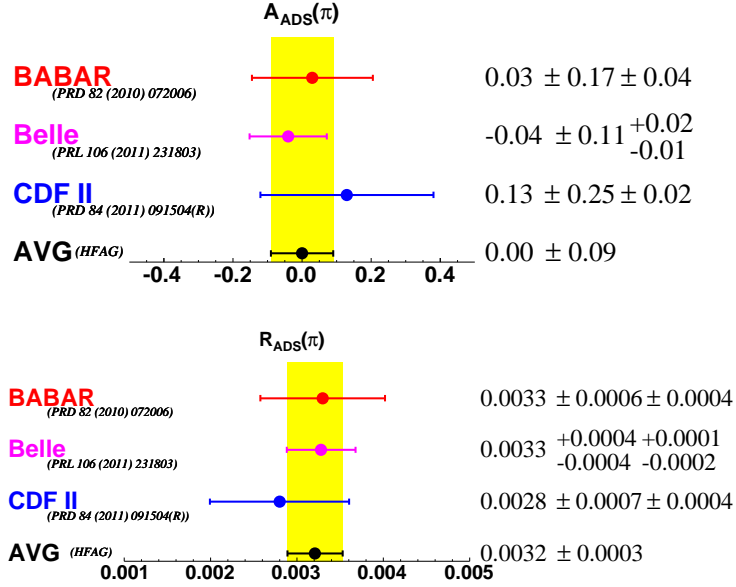

Figure 7.2: Comparison of $BABAR$, Belle and CDF experimental results on $A_{ADS}(\pi)$ and $R_{ADS}(\pi)$.
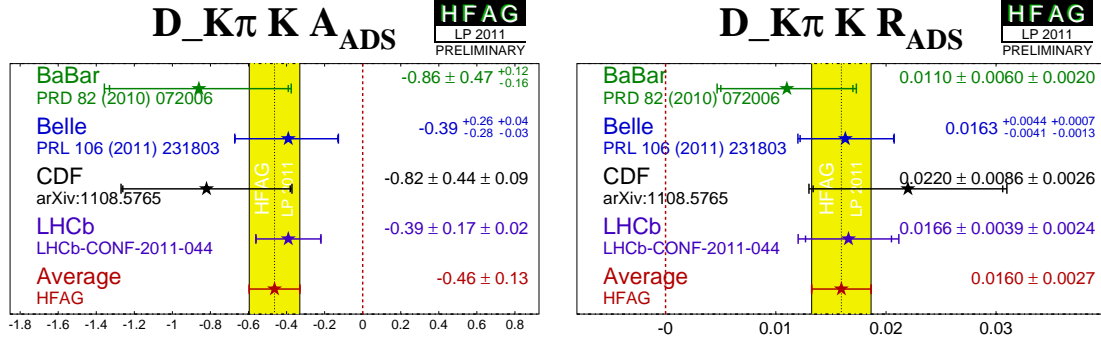
Figure 7.3: Comparison of $BABAR$, Belle, CDF and LHCb experimental results on $A_{ADS}(K)$ and $R_{ADS}(K)$.
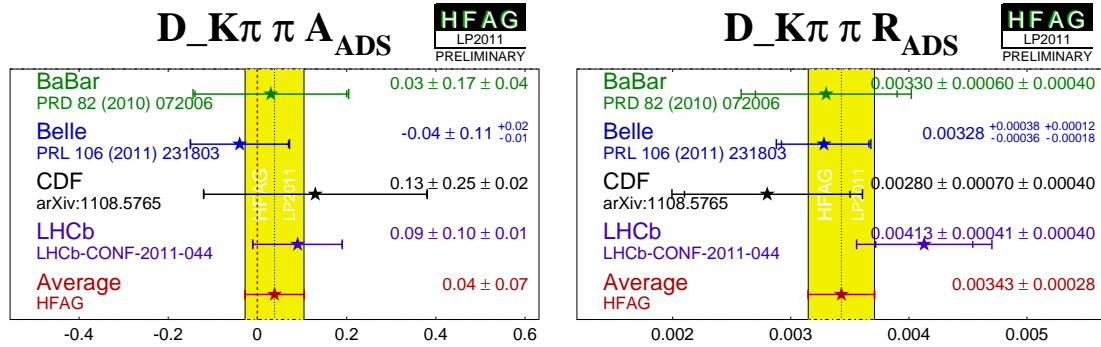

Figure 7.4: Comparison of $BABAR$, Belle, CDF and LHCb experimental results on $A_{ADS}(\pi)$ and $R_{ADS}(\pi)$.

Bibliography

[1] D. ATWOOD, I. DUNIETZ, and A. SONI, "Enhanced CP Violation with $B \rightarrow KD^0(\overline{D}^0)$ Modes and Extraction of the Cabibbo-Kobayashi-Maskawa Angle γ," *Phys. Rev. Lett.*, Apr 1997, vol. 78, pp. 3257–3260. http://link.aps.org/doi/10.1103/PhysRevLett.78.3257

[2] D. ATWOOD, I. DUNIETZ, and A. SONI, "Improved methods for observing CP violation in $B^{\pm} \rightarrow KD$ and measuring the CKM phase γ," *Phys. Rev. D*, Jan 2001, vol. 63, p. 036005. http://link.aps.org/doi/10.1103/PhysRevD.63.036005

[3] S. L. GLASHOW, "Partial Symmetries of Weak Interactions," *Nucl. Phys.*, 1961, vol. 22, pp. 579–588.

[4] A. SALAM, "Elementary particle physics." Almquist and Wiksell, 1968.

[5] S. WEINBERG, "A Model of Leptons," *Phys. Rev. Lett.*, 1967, vol. 19, pp. 1264–1266.

[6] D. GRIFFITHS, "Introduction to elementary particle." Wiley, 2008.

[7] F. HALZEN and A. D. MARTIN, "Quarks and leptons: an introductory course in modern particle physics." Wiley, 1984.

[8] M. E. PESKIN and V. SCHROEDER, "An introduction to quantum field theory." Westview Press, 1995.

[9] Q. HO-KIM and X. Y. PHAM, "Elementary particles and their interactions." Springer, 1998.

[10] I. I. BIGI and A. I. SANDA, "*CP* Violation." Cambridge University Press, 1999.

[11] M. SOZZI, "*CP* Violation." Oxford U.P., 2008.

[12] N. CABIBBO, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.*, 1963, vol. 10, pp. 531–533. http://link.aps.org/doi/10.1103/PhysRevLett.10.531

[13] M. KOBAYASHI and T. MASKAWA, "*CP*-Violation in the Renormalizable Theory of Weak Interaction," *Progress of Theoretical Physics*, 1973, vol. 49, no. 2, pp. 652–657. http://ptp.ipap.jp/link?PTP/49/652/

[14] L.-L. CHAU and W.-Y. KEUNG, "Comments on the Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.*, Nov 1984, vol. 53, pp. 1802–1805. http://link.aps.org/doi/10.1103/PhysRevLett.53.1802

[15] L. WOLFENSTEIN, "Parametrization of the Kobayashi-Maskawa matrix," *Phys. Rev. Lett.*, Nov 1983, vol. 51, pp. 1945–1947. http://link.aps.org/doi/10.1103/PhysRevLett.51.1945

[16] K. NAKAMURA, Particle Data Group Collaboration, "Review of particle physics," *J. Phys.*, 2010, vol. G37, p. 075021. http://dx.doi.org/10.1088/0954-3899/37/7A/075021

[17] C. JARLSKOG, "*CP* violation." World Scientific, 1989.

[18] R. ALEKSAN, B. KAYSER, and D. LONDON, "Determining the Quark Mixing Matrix from *CP*-violating Asymmetries," *Phys. Rev. Lett.*, Jul 1994, vol. 73, pp. 18–20. http://link.aps.org/doi/10.1103/PhysRevLett.73.18

[19] J. CHARLES ET AL., CKMfitter Group Collaboration, "*CP* Violation and the CKM Matrix Assessing the Impact of the Asymmetric B factories," *Eur.Phys.Jour.*, 2005, vol. C41, pp. 1–131, `arXiv:0406184 [hep-ex]`. http://ckmfitter.in2p3.fr

[20] M. BONA et al. UTfit Collaboration, "The UTfit collaboration report on the status of the unitarity triangle beyond the standard model. I: Model- independent analysis and minimal flavour violation," *JHEP*, 2006, vol. 03, p. 080, `arXiv:hep-ph/0509219`.

[21] D. ASNER ET AL., Heavy Flavor Averaging Group (HFAG) Collaboration, "Averages of b-hadron, c-hadron, and tau-lepton Properties," 2010," `arXiv:1010.1589 [hep-ex]`.

[22] P. F. HARRISON et al. BABAR Collaboration, "The BABAR physics book: Physics at an asymmetric B factory," papers from Workshop on Physics at an Asymmetric B Factory (BABAR Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[23] M. GRONAU and D. WYLER, "On determining a weak phase from charged b decay asymmetries," *Physics Letters B*, 1991, vol. 265, no. 1-2, pp. 172 – 176. http://www.sciencedirect.com/science/article/pii/037026939190034N

[24] M. GRONAU and D. LONDON, "How to determine all the angles of the unitarity triangle from $B_d^0 \to DK_s$ and $B_s^0 \to D\phi$," *Physics Letters B*, 1991, vol. 253, no. 3 - 4, pp. 483 – 488. http://www.sciencedirect.com/science/article/pii/037026939191756L

[25] A. GIRI, Y. GROSSMAN, A. SOFFER, and J. ZUPAN, "Determining γ using $B^\pm \to DK^\pm$ with multibody D decays," *Phys. Rev. D*, Sep 2003, vol. 68, p. 054018. http://link.aps.org/doi/10.1103/PhysRevD.68.054018

[26] M. GRONAU, "Weak phase γ from color-allowed $B \to DK$ rates," *Phys. Rev. D*, Jun 1998, vol. 58, p. 037301. http://link.aps.org/doi/10.1103/PhysRevD.58.037301

[27] A. BONDAR and T. GERSHON, "On ϕ_3 measurements using $B^- \to D^*K^-$ decays," *Phys. Rev. D*, Nov 2004, vol. 70, p. 091503. http://link.aps.org/doi/10.1103/PhysRevD.70.091503

[28] del Amo Sanchez et al. BaBar Collaboration, "Search for $b \rightarrow u$ transitions in $B^- \rightarrow DK^-$ and D^*K^- decays," *Phys. Rev. D*, Oct 2010, vol. 82, p. 072006. http://link.aps.org/doi/10.1103/PhysRevD.82.072006

[29] Y. Horii et al. Belle Collaboration, "Evidence for the Suppressed Decay $B^- \rightarrow DK^-$, $D \rightarrow K^+\pi^-$," *Phys. Rev. Lett.*, Jun 2011, vol. 106, p. 231803. http://link.aps.org/doi/10.1103/PhysRevLett.106.231803

[30] LHCb Collaboration, "Evidence for the suppressed decay $B^\pm \rightarrow (K^\mp\pi^\pm)DK^\pm$," Submitted to 2011 Europhysics Conference On High Energy Physics, Grenoble, France, 21 - 27 Jul 2011, CERN-LHCb-CONF-2011-044.

[31] G. Aubrecht, "A teachers guide to the nuclear science wall chart." Contemporary Physics Education Project, 2003.

[32] M. Popovic, L. Allen, and C. Schmidt, "Fermilab linac injector, revisited," prepared for 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High-energy Accelerators (IUPAP), Dallas, Texas, 1-5 May 1995.

[33] Fermilab Beam Division, *RUN II Handbook*, 1998. http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html

[34] D. Mohl, "Physics and technique of stochastic cooling," *Phys. Reports*, 1980, vol. 58.

[35] S. Nagaitsev et al., "Experimental demonstration of relativistic electron cooling," *Phys. Rev. Lett.*, 2006, vol. 96.

[36] D. Tonelli, "First observation of the $B_s^0 \rightarrow K^+K^-$ decay mode, and measurement of the B^0 and B_s^0 mesons decay-rates into two-body, charmless final states at CDF," Ph.D. dissertation, Scuola Normale Superiore di Pisa, 2006.

[37] CDF Collaboration, "The CDF IIb detector: Technical Design Report," FERMILAB-TM-2198.

[38] CDF COLLABORATION, "The CDF II Detector Technical Design Report," 1996, FERMILAB-PUB-96/390-E.

[39] CDF COLLABORATION, "Proposal for Ehnancement of the CDF II Detector: an Inner Silicon Layer and a Time of Flight Detector," 1998, FERMILAB-PROPOSAL-909 (and updates).

[40] D. ACOSTA et al. CDF Collaboration, "The performance of the CDF luminosity monitor," *Nucl. Instrum. Meth.*, 2002, vol. A494, pp. 57–62.

[41] F. ABE et al. CDF Collaboration, "The CDF detector: an overview," *Nucl. Instr. Meth.*, 1988, vol. A271, pp. 387–403, FERMILAB-PUB-94/024-E (1994).

[42] R. BLAIR et al. CDF Collaboration, "The CDF-II detector: Technical Design Report," FERMILAB-PUB-96/390-E (1996).

[43] T. NELSON CDF Collaboration, "The CDF Layer ØØ detector," *Int. J. Mod. Phys.*, 2001, vol. A16S1C, pp. 1091–1093.

[44] A. SILL CDF Collaboration, "CDF run II silicon tracking projects," *Nucl. Instrum. Meth.*, 2000, vol. A447, pp. 1–8.

[45] A. A. AFFOLDER et al., "Status report of the intermediate silicon layers detector at CDFII," *Nucl. Instrum. Meth.*, 2002, vol. A485, pp. 6–9.

[46] K. T. PITTS et al. CDF Collaboration, "The CDF central outer tracker," *Nucl. Phys. Proc. Suppl.*, 1998, vol. 61B, pp. 230–235.

[47] C. S. HILL et al. CDF Collaboration, "Operational experience and performance of the CDFII silicon detector," *Nucl. Instrum. Meth.*, 2004, vol. A530, pp. 1–6.

[48] http://penn01.fnal.gov

[49] S. D'AURIA, D. LUCCHESI, S. D. RONCO, R. CAROSI, M. CIOCCI, S. DONATI, P. CATASTINI, M. MORELLO, G. PUNZI, P. SQUILLACIOTI, D. TONELLI, S. TORRE, M. RESCIGNO, S. D. CECCO, S. GIAGU, M. CASARSA, and A. CERRI, "Track-based calibration of the COT specific ionization," 2004, CDF/ANAL/BOTTOM/CDFR/6932.

[50] S. E. Yu, "COT dE/dx measurement and corrections," 2004, CDF/DOC/BOTTOM/PUBLIC/6361.

[51] P. Azzi, G. Busetto, P. Gatti, and A. Ribon, "Histogram Tracking in the COT," 2001, CDF/DOC/TRACKING/CDFR/5562.

[52] F. Snider et al. CDF Collaboration, "Tracking at CDF: Algorithms and experience from Run I and Run II," *Nucl. Instrum. Meth.*, 2006, vol. A566, pp. 133–141.

[53] C. Paulk et al. CDF Collaboration, "Inside-out tracking at CDF," *Nucl. Instrum. Meth.*, 2005, vol. A538, pp. 249–254.

[54] J. Goldstein, C. Issever, T. Nelson, R. Snider, and D. Stuart, "Silicon tracking for plug electrons," 2002, CDF/DOC/TRACKING/CDFR/5970.

[55] D. Acosta et al. CDF Collaboration, "A time-of-flight detector in CDF-II," *Nucl. Instrum. Meth.*, 2004, vol. A518, pp. 605–608.

[56] C. Ginsburg et al. CDF Collaboration, "CDF run 2 muon system," *Eur. Phys. J.*, 2004, vol. C33, pp. s1002–s1004.

[57] G. Ascoli et al., "CDF central muon detector," *Nucl. Instrum. Meth.*, 1988, vol. A268, p. 33.

[58] S. Klimenko, J. Konigsberg, and T. M. Liss, "Averaging of the inelastic cross sections measured by the CDF and the E811 experiments," 2003, FERMILAB-FN-0741.

[59] G. Gomez-Ceballos et al., "Event builder and Level 3 at the CDF experiment," *Nucl. Instrum. Meth.*, 2004, vol. A518, pp. 522–524.

[60] The Trigger and Datasets Working Group, "Run II trigger table and datasets plan," 2001, CDF/PHYS/TRIGGER/CDFR/4718.

[61] S. R. Amendolia et al., "The AMchip: a Full-custom CMOS VLSI Associative Memory for Pattern Recognition," *IEEE Trans. Nucl. Sci.*, 1992, vol. 39, pp. 795–797.

[62] W. ASHMANSKAS et al. CDF Collaboration, "Performance of the CDF online silicon vertex tracker," *IEEE Trans. Nucl. Sci.*, 2002, vol. 49, pp. 1177–1184.

[63] R. BRUN et al., *GEANT: Simulation Program For Particle Physics Experiments. User Guide and Reference Manual*, 1978.

[64] J. MARRINER, "Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT)," 1993, Tech. Rep. CDF/DOC/SEC VTX/PUBLIC/1996.

[65] G. PUNZI, "Sensitivity of Searches for New Signals and Its Optimization," *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, 2003, p. 79, Proc. of the Conference PhyStat 2003, `arXiv:physics/0308063 [data-an]`.

[66] J. D. BJORKEN, "Properties of hadron distributions in reactions containing very heavy quarks," *Phys. Rev. D*, Jan 1978, vol. 17, pp. 171–173. http://link.aps.org/doi/10.1103/PhysRevD.17.171

[67] G. PUNZI, "Notes on statistical separation of classes of events," Nov. 2006, CDF/MEMO/STATISTICS/PUBLIC/8284, `arXiv:physics/0611219 [data-an]`.

[68] P. MARINO, P. GAROSI, and G. PUNZI, "Study of $B_s^0 \to D_s\pi$ and $B_s^0 \to D_s K$ decays with a multivariate approach," 2011, CDF/PHYS/BOTTOM/GROUP/10772.

[69] E. BARACCHINI and G. ISIDORI, "Electromagnetic corrections to non-leptonic two-body b and d decays," *Physics Letters B*, 2006, vol. 633, no. 2-3, pp. 309 – 313. http://www.sciencedirect.com/science/article/pii/S0370269305017284

[70] G. NANAVA and Z. WAS, "Scalar QED, NLO and PHOTOS Monte Carlo," *The European Physical Journal C - Particles and Fields*," vol. 51, pp. 569–583. http://dx.doi.org/10.1140/epjc/s10052-007-0316-5

[71] P. GOLONKA and Z. WAS, "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays," *Eur. Phys. J.*, 2006, vol. C45, pp. 97–107, `arXiv:hep-ph/0506026`.

[72] P. GAROSI, P. SQUILLACIOTI, and G. PUNZI, "Measurements of Branching Fractions and CP Asymmetries in the DCS decay modes $B^- \to Dh^-$ on 7 fb^{-1}," 2011, CDF/PHYS/BOTTOM/GROUP/10264.

[73] G. COWAN, "Statistical Data Analysis." Clarendon Press, 1998.

[74] F. JAMES and M. ROOS, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.*, 1975, vol. 10, pp. 343–367.

[75] F. JAMES, "MINUIT - Function Minimization and Error Analysis - Reference Manual," 1998, CERN Program Library D506.

[76] G. PUNZI, "Comments on Likelihood fits with variable resolution," Jan. 2004," `arXiv:physics/0401045 [data-an]`.

[77] D. ACOSTA et al. CDF Collaboration, "Measurement of Partial Widths and Search for Direct CP Violation in D^0 Meson Decays to $K^- K^+$ and $\pi^- \pi^+$," *Phys. Rev. Lett.*, Apr 2005, vol. 94, p. 122001. http://link.aps.org/doi/10.1103/PhysRevLett.94.122001

[78] T. AALTONEN et al. CDF Collaboration, "Measurements of branching fraction ratios and CP asymmetries in $B^\pm \to D_{CP}K^\pm$ decays in hadron collisions," *Phys. Rev. D*, Feb 2010, vol. 81, p. 031105. http://link.aps.org/doi/10.1103/PhysRevD.81.031105

[79] T. AALTONEN et al. CDF Collaboration, "Measurements of Direct CP Violating Asymmetries in Charmless Decays of Strange Bottom Mesons and Bottom Baryons," *Phys. Rev. Lett.*, May 2011, vol. 106, p. 181802. http://link.aps.org/doi/10.1103/PhysRevLett.106.181802